

10/07/09

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 09:14:16 ; Search time 6001 Seconds
(without alignments)
11634.333 Million cell updates/sec

Title: AJ012098

Perfect score: 2399
Sequence: 1 ATCTTACATGAAACACACAAA.....TGTAAACAGCAGCACTTCTGAG 2399

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_ses:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pal:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_ses:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_in:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_higo_hum:*
- 40: em_higo_mus:*
- 41: em_higo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2399	100.0	2399	8	CRE012098	AJ012098 Chlamydom
2	2218	92.5	2522	8	AF289201	AF289201 Chlamydom
3	2119.4	88.3	2454	8	AY055755	AY055755 Chlamydom
4	799.4	33.1	2560	8	AY055756	AY055756 Chlamydom
5	448.6	18.7	2420	8	CF0298228	CF0298228 Chlorella
6	424.6	11.7	1607	8	AF276706	AF276706 Scenedesmus
7	286.2	11.9	5251	1	DFU07229	DFU07229 Desulfococcus
8	244.6	10.2	5001	8	SOB271546	SOB271546 Scenedesmus
9	242.4	10.1	2887	1	DFHYDABG	DFHYDABG D. vulgaris
10	222	9.3	1964	1	DVHYDPE	DVHYDPE D. vulgaris
11	201.4	8.4	2067	1	DVHYDG	DVHYDG D. vulgaris
12	186	7.8	3265	8	CFU298227	CFU298227 Chlorella
13	180.8	7.5	2585	1	AF120457	AF120457 Megasthrix
14	166.6	6.9	15158	1	AE001705	AE001705 Thermotoga
15	160	6.7	1940	1	DVHYDC	DVHYDC D. vulgaris
16	159.4	6.6	1800	1	AF331719	AF331719 Desulfococcus
17	148.2	6.2	11577	1	AE013056	AE013056 Thermotoga
18	144.8	6.0	1848	3	AF262401	AF262401 Trichomonas
19	134	5.6	16133	1	EAC312124	EAC312124 Eubacterium
20	124.4	5.2	2238	1	AF148212	AF148212 Clostridium
21	120.2	5.0	1109	3	AF446077	AF446077 Trichomonas
22	109.6	4.6	1109	3	AY028641	AY028641 Trichomonas
23	100.8	4.2	1659	3	TVU26964	TVU26964 Trichomonas
24	89.6	3.7	3625	3	NYHYDROG	NYHYDROG Nyctotherus
25	86.4	3.6	1544	3	TVU19897	TVU19897 Trichomonas
26	81.2	3.4	357	3	NOV420000	NOV420000 Nyctotherus
27	80.2	3.3	113193	1	AF357202	AF357202 Streptococcus
28	79.2	3.3	45313	1	SC95A	SC95A Streptococcus
29	78.2	3.3	34893	1	SC17	SC17 Streptococcus
30	78.2	3.2	125020	9	AF429315	AF429315 Homo sapi
31	77.8	3.2	125020	9	AF429315	AF429315 Homo sapi
32	77.6	3.2	1761	8	AF446076	AF446076 Pityrospora
33	76.8	3.2	3113	8	SS95BACAL	SS95BACAL Streptococcus
34	76.6	3.2	34611	1	SCD63A	SCD63A Streptococcus
35	75.8	3.2	33610	1	SCB6	SCB6 Streptococcus
36	75.4	3.1	2467	8	AY033895	AY033895 Neocallim
37	75	3.1	109519	6	AX195929	AX195929 Streptococcus
38	74	3.1	1389	1	SL441125	SL441125 Streptococcus
39	74	3.1	1728	6	AX196054	AX196054 Streptococcus
40	74	3.1	42210	1	SC1C2	SC1C2 Streptococcus
41	73.6	3.1	30253	1	SCB84	SCB84 Streptococcus
42	73.4	3.1	1109	3	AY028640	AY028640 Trichomonas
43	73.4	3.1	12097	1	AE005116	AE005116 Halobacter
44	73	3.0	22991	1	SCD19	SCD19 Streptococcus
45	71.8	3.0	25362	1	SCF56	SCF56 Streptococcus

ALIGNMENTS

RESULT 1	CRE012098	2399 bp	mRNA	PLN 15-NOV-2001
LOCUS	CRE012098			
DEFINITION	Chlamydomonas reinhardtii mRNA for Fe-hydrogenase.			
ACCESSION	AJ012098			
VERSION	AJ012098.1 GI:16945125			
KEYWORDS	Fe-hydrogenase; hsdA gene.			
SOURCE	Chlamydomonas reinhardtii.			
ORGANISM	Chlamydomonas reinhardtii			
REFERENCE	1	Kaminski, A.U. and Happe, T.		
AUTHORS		Isolation and characterization of the hsdA gene encoding the		
TITLE		Fe-hydrogenase of Chlamydomonas reinhardtii		

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2399)
AUTHORS Happe, T.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-1998) Happe T., Molecular Biochemistry, Botanical
Institute, Kirchallee 1, D53115-Bonn, GERMANY
REMARK Revised by author 31-JAN-2001
COMMENT Related sequence AJ308413.
FEATURES
Source Location/Qualifiers
1..2399
/organism="Chlamydomonas reinhardtii"
/db_xref="taxon:3055"
141..1652
/gene="hyda"
/gene="hyda"
141..308
/gene="hyda"
159..1652
/gene="hyda"
/EC_number="1.18.99.1"
/function="hydrogen evolution"
/codon_start=1
/evidence="experimental"
/product="Fe-hydrogenase"
/protein_id="CAC80065.1"
/db_xref="GI:16945126"
/translation="MSALVTKPCAASVIRSSCRAROVAPRAPLASTRVALATLEA
PARRLGNVACAAAPAPAEAPLPSHQALSLAPKDDPTKRVCAVAPAVRALIET
LGLAPGATTKQALBEGRLRGLPDEVEFDLFGADLTMEBSSELHLRLEHLEAPSD
EPLPMFTSCCPGWIMALEKSPDLIPVSSCKSPOMMLAMVKSYLEKKGIAPKDMV
MSVIMPTCKROSEADRDWCVDADPLRLQDHYITVEIGNIPEKGINIAELPEDEM
DNPMVGSGAGVLEFTGGVMEALRPAFLFTGPPLPRLSIEVANGINAEIKETNTM
VPAAGSFEELHGRBAAPRAEAAGHTPGVPLADGAGATSEBGRGILLRAVANIET
GNAKKILTKQABAKYDPEYINACPAAGVGGGQPRSDKATITQRQALVNLDEKS
TKRSHENBSIRLELYDYLGEPLGKHAHELTHYVAGVEKDEKK"
309..1640
/gene="hyda"
/product="Fe-hydrogenase"
BASE COUNT 469 a 745 c 765 g 420 t
ORIGIN
Query Match 100.0%; Score 2399; DB 8; Length 2399;
Best Local Similarity 100.0%; Pred. No. 2,8e-287;
Matches 2399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCTTACATGACACCAAACTCTCGACGACCTCAAGCTCAAACTCGAAACCTTTT 60
DB 1 ATCTTACATGACACCAAACTCTCGACGACCTCAAGCTCAAACTCGAAACCTTTT 60
QY 61 TCCACAGTTTACACCCCAATTGAGACGCGCTCAAGCTCGCTCGCTTGTCTTCTATC 120
DB 61 TCCACAGTTTACACCCCAATTGAGACGCGCTCAAGCTCGCTCGCTTGTCTTCTATC 120
QY 121 GCACCACTTATTTCTTAATATGTAAGACGACCAAGATGTCGGCTGTGTGTAAGC 180
DB 121 GCACCACTTATTTCTTAATATGTAAGACGACCAAGATGTCGGCTGTGTGTAAGC 180
QY 181 CCTGCGCGCGCGTGTATTCGCGGAGCTCTGACGGCGCGCGAGTCGCCCCCGCG 240
DB 181 CCTGCGCGCGCGTGTATTCGCGGAGCTCTGACGGCGCGCGAGTCGCCCCCGCG 240
QY 241 CTCGCGTCCGAGCAGACCGTGTAGCCCTTTCACCACTTGAAGCGCCGACGCG 300
DB 241 CTCGCGTCCGAGCAGACCGTGTAGCCCTTTCACCACTTGAAGCGCCGACGCG 300
QY 301 GCTTAGGCAACGTCGCTTGGCGCGCTGCGACCCGCTCGAGGCGCTTTGAGTCATG 360
DB 301 GCTTAGGCAACGTCGCTTGGCGCGCTGCGACCCGCTCGAGGCGCTTTGAGTCATG 360
QY 361 TCCAGAGGCGCTCGCGAGCTTGGCAAGCCCAAGAGACGACCCCAAGCGCAAGCTGT 420
DB 361 TCCAGAGGCGCTCGCGAGCTTGGCAAGCCCAAGAGACGACCCCAAGCGCAAGCTGT 420
QY 421 GCGTGCAGGTGCGTCCGCGCTTGTGTGCTATTGCGAGACCCCTGGGCTTGGCGCGG 480

DB 421 GCGTGCAGGTGCGTCCGCGCTTGTGTGCTATTGCGAGACCCCTGGGCTTGGCGCGG 480
QY 481 GCGGCACACCCCAAGACGCTTGGCCGAGAGGCTCCGCGCTTGGAGAGTGT 540
DB 481 GCGGCACACCCCAAGACGCTTGGCCGAGAGGCTCCGCGCTTGGAGAGTGT 540
QY 541 TTGACACGCTGTTTGGCGCGCAGCTTGAACATCATGAGAGAGGACGAGCTGTGACC 600
DB 541 TTGACACGCTGTTTGGCGCGCAGCTTGAACATCATGAGAGAGGACGAGCTGTGACC 600
QY 601 GCTTCCAGACGACCTTGAAGGCCCCCACTCCGACGAGCGCTGCTTGTACCA 660
DB 601 GCTTCCAGACGACCTTGAAGGCCCCCACTCCGACGAGCGCTGCTTGTACCA 660
QY 661 GCTGCTGCCCGCGCTGATCGCTATGCTGTGAGAAATCTTACCAGGCTGATCCCTAG 720
DB 661 GCTGCTGCCCGCGCTGATCGCTATGCTGTGAGAAATCTTACCAGGCTGATCCCTAG 720
QY 721 TGAGCAGCTGCAAGAGCCCCCAGATGATGCTGGCGCAATGCTCAAGTCTTACTAGCGG 780
DB 721 TGAGCAGCTGCAAGAGCCCCCAGATGATGCTGGCGCAATGCTCAAGTCTTACTAGCGG 780
QY 781 AAAAGAGGGCATTCGCGCCCAAGAGCATGCTATGCTTCATCATGCCCTTCACGCGCA 840
DB 781 AAAAGAGGGCATTCGCGCCCAAGAGCATGCTATGCTTCATCATGCCCTTCACGCGCA 840
QY 841 ACCAGTCGAGGCTGACCGCGACTGCTGTGTGAGACGCCACCCACCTCGCGCGAC 900
DB 841 ACCAGTCGAGGCTGACCGCGACTGCTGTGTGAGACGCCACCCACCTCGCGCGAC 900
QY 901 TGACCACTGTCATCAACACCGTGTGAGCTGGGCAATCTTCAAGAGCGCGCATCAAC 960
DB 901 TGACCACTGTCATCAACACCGTGTGAGCTGGGCAATCTTCAAGAGCGCGCATCAAC 960
QY 961 TGGCGGAGCTGCGCGAGGGCGAGTGGACAAATCCAAATGGCGCTGGCGCGCGCG 1020
DB 961 TGGCGGAGCTGCGCGAGGGCGAGTGGACAAATCCAAATGGCGCTGGCGCGCGCG 1020
QY 1021 TGCTGTTGGGACCAACCGCGGTGTCTATGAGAGCGCGCTGCGACAGGCTTATGAGCTGT 1080
DB 1021 TGCTGTTGGGACCAACCGCGGTGTCTATGAGAGCGCGCTGCGACAGGCTTATGAGCTGT 1080
QY 1081 TCAAGGCGACCGCGCTGCGCGCTGAGCTTGAAGAGTGTGCGGCAATGACATCA 1140
DB 1081 TCAAGGCGACCGCGCTGCGCGCTGAGCTTGAAGAGTGTGCGGCAATGACATCA 1140
QY 1141 AGGAGACCAATCACTACATGATGCGCGCGCGCTTCAAGTTTGAAGAGCTGTGAAGC 1200
DB 1141 AGGAGACCAATCACTACATGATGCGCGCGCGCTTCAAGTTTGAAGAGCTGTGAAGC 1200
QY 1201 ACCGCGCGCGCGCGCGCGCGAGGCGCGCGCGCAAGCCCGCGCGCTGCGCTGGG 1260
DB 1201 ACCGCGCGCGCGCGCGCGCGAGGCGCGCGCGCAAGCCCGCGCGCTGCGCTGGG 1260
QY 1261 ACGGCGCGCGCGCTTCAACAGAGAGCGAGGCGCGCATCATCTGCGCGTGGCGG 1320
DB 1261 ACGGCGCGCGCGCTTCAACAGAGAGCGAGGCGCGCATCATCTGCGCGTGGCGG 1320
QY 1321 TGGCCAAAGGCGTGGGCAAGCGCAAGAGCTGATCACCAAGATGAGGCGCGCGAGCA 1380
DB 1321 TGGCCAAAGGCGTGGGCAAGCGCAAGAGCTGATCACCAAGATGAGGCGCGCGAGCA 1380
QY 1381 AGTACGACTTTGTGAGATCATGAGCTTGCCTCGCGCGCTGTGTGAGCGCGCGCGCAGC 1440
DB 1381 AGTACGACTTTGTGAGATCATGAGCTTGCCTCGCGCGCTGTGTGAGCGCGCGCGCAGC 1440
QY 1441 CCCGCTTCCACCAAGGCGCATCAAGAGAGCGGCGCGCTGTACAACCTTGAAGC 1500
DB 1441 CCCGCTTCCACCAAGGCGCATCAAGAGAGCGGCGCGCTGTACAACCTTGAAGC 1500
QY 1501 AGAAGTCCACGCTGCGCGCGAGCGAGCAAGCCGCTCATCGGAGCTGTACGACAGT 1560

Db 1501 AGAATCCACGCTGCCCGCAGCCACGAGAACCCGTCATCCCGAGCTGTACGACCGT 1560
QY 1561 ACCTCCGAGAGCGCTGTGGCCACAGAGCGCAGAGCTGTGACACCCACTAGTGGCCG 1620
Db 1561 ACCTCCGAGAGCGCTGTGGCCACAGAGCGCAGAGCTGTGACACCCACTAGTGGCCG 1620
QY 1621 GCGGCGTGAAG 1680
Db 1621 GCGGCGTGAAG 1680
QY 1681 CAGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
Db 1681 CAGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
QY 1741 GGTGTCTCTCTTATATATGAGTGGGTCAGAAATAGCGGGGCTCAAAATCTTTCTTTT 1800
Db 1741 GGTGTCTCTCTTATATGAGTGGGTCAGAAATAGCGGGGCTCAAAATCTTTCTTTT 1800
QY 1801 GAGTGTGTACAGCATGGGGGACGTGTGCGAGGCGAGTGGCTGTTCACGACGCTG 1860
Db 1801 GAGTGTGTACAGCATGGGGGACGTGTGCGAGGCGAGTGGCTGTTCACGACGCTG 1860
QY 1861 GCAATTAGGGGTAGTACTGTGCAATGAGGAGCGGCGCTGTGTAACGAAATGGCGTATCCCT 1920
Db 1861 GCAATTAGGGGTAGTACTGTGCAATGAGGAGCGGCGCTGTGTAACGAAATGGCGTATCCCT 1920
QY 1921 CCAGGAGCAGCTCGGAATGGCGCGCTGCCATCAACGCAAAATCTTGGCTTCATCGCTCT 1980
Db 1921 CCAGGAGCAGCTCGGAATGGCGCGCTGCCATCAACGCAAAATCTTGGCTTCATCGCTCT 1980
QY 1981 GGAATATGAGCTGACAAACCTGCAATTCATTTGCTTTTACAGCTGCCCAATCTTG 2040
Db 1981 GGAATATGAGCTGACAAACCTGCAATTCATTTGCTTTTACAGCTGCCCAATCTTG 2040
QY 2041 GTTGGAGAGCTAAACATGTTTGGGACAACTTAACTTAACTTAACTTAACTTAACTTAA 2100
Db 2041 GTTGGAGAGCTAAACATGTTTGGGACAACTTAACTTAACTTAACTTAACTTAACTTAA 2100
QY 2101 TGGGACGTTGTGCGCTGTGAGTGGGCGGGAACGTGGGTAGGATTTAGGCTGTGGCA 2160
Db 2101 TGGGACGTTGTGCGCTGTGAGTGGGCGGGAACGTGGGTAGGATTTAGGCTGTGGCA 2160
QY 2161 TACGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
Db 2161 TACGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
QY 2221 TTATACCAACCTGATGAGTGAAGCTTGTGTAGGCAACGAGCGGTAGGAGGAGGAGGAG 2280
Db 2221 TTATACCAACCTGATGAGTGAAGCTTGTGTAGGCAACGAGCGGTAGGAGGAGGAGGAG 2280
QY 2281 GATGCAATTGCAACGCTGTAAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
Db 2281 GATGCAATTGCAACGCTGTAAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
QY 2341 TTGGCCGAGCAG 2399
Db 2341 TTGGCCGAGCAG 2399

TITLE The iron hydrogenase of *Chlamydomonas reinhardtii* has a single folding domain containing an H-cluster catalytic center and lacking bound electron carriers
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2522)
AUTHORS Metz, L.J.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-2000) Molecular Genetics and Cell Biology, The University of Chicago, 1103 E. 57th St., Chicago, IL 60637, USA
FEATURES
source 1..2522
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/map="LG_III"
/clone="CLS8b09"
1..2522
/gene="hyd1"
135..1628
/gene="hyd1"
/EC number="1.18.99.1"
/note="Method: conceptual translation with partial peptide sequencing"
/codon_start=1
/product="Fe-hydrogenase precursor"
/protein_id="AAG00591.1"
/db_xref="GI:9837540"
/translation="MSALVLRKPAVSIIRGSSCRARQVAPAPLAATRVATLEA
PARRIGNVACAAAPAEAPLHVQALDELKPKDDPRKRVQVAAVRYALLET
LGLAPGATPKQLAGLRLGFDEVFDLFGADLTIMESSELHRLLEAPHSID
EPLPMTSCCPGMIMLEKSYDPLPYVSSCSQPMMLAMVKSILAEKXGIALPKDMV
AVSIMPTCKROSEADRMFCVDADPLRLDHLVITVELGNIFKEKGINLAEPDEM
DNMPGVSGAGVLPFTGGVMEALRTALVLPFTGPIPLSLSEVARGMURKENTITM
VAPRSKFEELIKHRAARAAEAAAGTPEPLAMDGAGCTSDRGITLRVAVANGL
GNAKULITMQAGEAKTIDPEITMACPAGCVGGGQPRSDKAITQRQALYNLDEKS
TLRSHENPSIRELYDTYLGEPHKAHELTHYVAVGVEKDEKK"
transit_peptide 135..302
/gene="hyd1"
303..1625
/gene="hyd1"
mat_peptide /product="Fe-hydrogenase"
BASE COUNT 484 a 777 c 812 g 449 t
ORIGIN
Query Match 92.5%; Score 2218; DB 8; Length 2522;
Best Local Similarity 94.2%; Pred. No. 6, 2e-265;
Matches 2375; Conservative 0; Mismatches 0; Indels 147; Gaps 1;
QY 25 CTGCGAGGACTAGCTCAAAACCTTCTTCCAGAGTTTACACCCCAATTGG 84
Db 1 CTGCGAGGACTAGCTCAAAACCTTCTTCCAGAGTTTACACCCCAATTGG 60
QY 85 GAGCGGCTCAAGCTGCTCGGTGCTCTTCATGCAACCACTTATTTCTAATATC 144
Db 61 GAGCGGCTCAAGCTGCTCGGTGCTCTTCATGCAACCACTTATTTCTAATATC 120
QY 145 GTAGAGCGCAAGAGTGTGGGCTGTGTGCTGAAGCCGCGCGCGGTCTATTCGGG 204
Db 121 GTAGAGCGCAAGAGTGTGGGCTGTGTGCTGAAGCCGCGCGCGGTCTATTCGGG 180
QY 205 GCAGCTCTGACAGGAGGCGGAGGTGCGCCCGCGCTCGCTGCGAGCAGACCGGTGC 264
Db 181 GCAGCTCTGACAGGAGGCGGAGGTGCGCCCGCGCTCGCTGCGAGCAGACCGGTGC 240
QY 265 GTGTAGCCCTTGCAACATTGAGGCGCCGCAAGCGGCTTAGGCAAGTGTGCTGGCGG 324
Db 241 GTGTAGCCCTTGCAACATTGAGGCGCCGCAAGCGGCTTAGGCAAGTGTGCTGGCGG 300
QY 325 CTGCGGACCCGCTGGGAGGCGGCTTTAGTACATGTCCAGAGAGGCTGCGGACCTTG 384
Db 301 CTGCGGACCCGCTGGGAGGCGGCTTTAGTACATGTCCAGAGAGGCTGCGGACCTTG 360
QY 385 CCAGCCCAAGAGCAGACCCACGCGCAACAGTGTGAGTGGTGTCCGCGCGTTC 444

Db	361	CCAAAGCCCAAGAGACCCCAACCGGACAGACGCTTCGCGAGGTGGCTTCGGGCGCTTC	420
Qy	445	GTGTGCTATTGCCGAGACCTTGAGCTGGGCGCGGGCGCCACACCCCAAGACGTGG	504
Db	421	GTGTGCTATTGCGAGACCTTGAGCTGGGCGCGGGCGCCACACCCCAAGACGTGG	480
Qy	505	CCGAGGGGCTTCGGCGCGCTGGGGCTTGAAGAGGAGTTTGAACAGCGTTTGGCGCCAGC	564
Db	481	CCGAGGGGCTTCGGCGCGCTGGGGCTTGAAGAGGAGTTTGAACAGCGTTTGGCGCCAGC	540
Qy	565	TGACCATCATGAGAGAGAGGAGCAGAGCTGCTGACACCGCTTCACCGACACTTGAAGGCC	624
Db	541	TGACCATCATGAGAGAGGAGCAGAGCTGCTGACACCGCTTCACCGACACTTGAAGGCC	600
Qy	625	ACCCGCACTCCGAGAGCGGCTCCCATATGTTACACAGCTGCTGCCCCGGCTGGAATGGCTA	684
Db	601	ACCCGCACTCCGAGAGCGGCTCCCATATGTTACACAGCTGCTGCCCCGGCTGGAATGGCTA	660
Qy	685	TGCTGAGAAATCTTACCAGGACCTGATCCCTACGCTGAGACAGCTCAGAGGCCCCAGA	744
Db	661	TGCTGAGAAATCTTACCAGGACCTGATCCCTACGCTGAGACAGCTCAGAGGCCCCAGA	720
Qy	745	TGATGCTGGGGGCGCATGCTCAAGTCTTACTACCGGAAAAGAGGGCATTCGGCCAAAG	804
Db	721	TGATGCTGGGGGCGCATGCTCAAGTCTTACTACCGGAAAAGAGGGCATTCGGCCAAAG	780
Qy	805	ACATGCTCATGTGTCCATCATATGCTCTGACGAGCGGACAGTCCGAGGCTGAACCGGACT	864
Db	781	ACATGCTCATGTGTCCATCATATGCTCTGACGAGCGGACAGTCCGAGGCTGAACCGGACT	840
Qy	865	GGTTCTGTGTGACGCGCCACCCACCTTGCGCCAGCTGAGACCAAGTATACACACCGTGG	924
Db	841	GGTTCTGTGTGACGCGCCACCCACCTTGCGCCAGCTGAGACCAAGTATACACACCGTGG	900
Qy	925	AGCTGGGAAACATCTTCCAGAGCGGGGGAATCAACTGGGCGAGCTGCCGAGGGCGAGT	984
Db	901	AGCTGGGAAACATCTTCCAGAGCGGGGGAATCAACTGGGCGAGCTGCCGAGGGCGAGT	960
Qy	985	GGGACAAATCCAAATGGGCGTGGGCTCGGGCGCGGCGCTGCTGTTCGGACACACCGGCGTG	1044
Db	961	GGGACAAATCCAAATGGGCGTGGGCTCGGGCGCGGCGCTGCTGTTCGGACACACCGGCGTG	1020
Qy	1045	TGATGAGGCGGGGCTGCGCCAGGCGCTATGAGCTGTTACGGGACAGCGGCTGGCGCGC	1104
Db	1021	TGATGAGGCGGGGCTGCGCCAGGCGCTATGAGCTGTTACGGGACAGCGGCTGGCGCGC	1080
Qy	1105	TGAGCTGTGACGAGGTGCGCGGATGAGCGGATCAAGAGACCAATCATACATGAGTGG	1164
Db	1081	TGAGCTGTGACGAGGTGCGCGGATGAGCGGATCAAGAGACCAATCATACATGAGTGG	1140
Qy	1165	CCGCGCCCGGGTCCAAAGTTTGAAGAGCTGCTGAAGACACCGGCGCGCGCGCGCGAGG	1224
Db	1141	CCGCGCCCGGGTCCAAAGTTTGAAGAGCTGCTGAAGACACCGGCGCGCGCGCGCGAGG	1200
Qy	1225	CCGCGCGCGAGCGACCCCGGGGCGCTGAGCTTGAGACGCGCGGCGGGCTTACACAGG	1284
Db	1201	CCGCGCGCGAGCGACCCCGGGGCGCTGAGCTTGAGACGCGCGGCGGGCTTACACAGG	1260
Qy	1285	AGGACGGCAAGGAGCGGCATCACATGCGGTGAGGCGGCTGAGCCACCGGCTGGGCAACGGCA	1344
Db	1261	AGGACGGCAAGGAGCGGCATCACATGCGGTGAGGCGGCTGAGCCACCGGCTGGGCAACGGCA	1320
Qy	1345	AGAACTGTATCAACAAAGATGACAGGCGGGGAGAGGCCAAGTAGACACTTTGTGAGATCATAGG	1404
Db	1321	AGAACTGTATCAACAAAGATGACAGGCGGGGAGAGGCCAAGTAGACACTTTGTGAGATCATAGG	1380
Qy	1405	CTGTGCCCCGAGGCTGTGTGAGCGGCGGCGGCGAGCCCGCTTCACCGACGACAGGCGATCA	1464
Db	1381	CTGTGCCCCGAGGCTGTGTGAGCGGCGGCGGCGAGCCCGCTTCACCGACGACGCGATCA	1440
Qy	1465	CGCAGAAAGGGGAGGCGGCGCTGTACAACTTGACGAGAAAGTCAACGCTGCGCGCGACCC	1524

Db	1441	CGCAAGACGGCAGGCGGCGCTGTACACCTGACAGAAACTCCACCGTGGCCGACGC	1500
OY	1525	ACGAGAAACCCCTCCATCCCGGAGCTGTATGCACAGTACTTGGAGAGCCCTGGGCACA	1584
Db	1501	ACGAAACCCCGTCCATCCCGCAGCGCTGTATGCACAGTACTTGGAGAGCCCTGGGCACA	1566
OY	1585	AGGCCACAGAGCTGCTGCACACCCACTACGAGGCGGCGGCTGGAGAGAAAGACGAGA	1644
Db	1561	AGGCCACAGAGCTGCTGCACACCCACTACGAGGCGGCGGCTGGAGAGAAAGACGAGA	1620
OY	1645	AGAACTGAGAGCGCCACAGGCTCTTTGGCGGAGACAGCTTCAAAGCGAGGCGGTAT	1704
Db	1621	AGAACTGAGAGCGCCACAGGCTCTTTGGCGGAGACAGCTTCAAAGCGAGGCGGTAT	1680
OY	1705	TAGCAGTACCCATAATATGCACTGATGGGTATATGCGGGTGTCTCTTTATATTGAATGG	1764
Db	1681	TAGCAGTACCCATAATATGCACTGATGGGTATCGGGGTGTCTCTTTATATTGAATGG	1740
OY	1765	GATCAAAATAGCGCGCGGCTAAATGTCTTTTGTAGTGGTGTACAGCATGGGCAC	1824
Db	1741	GATCAAAATAGCGCGCGGCTAAATGTCTTTTGTAGTGGTGTACAGCATGGGCAC	1800
OY	1825	GTGTGCGGAGGCTAG-----	1839
Db	1801	GTGTGCGGAGGCTAGTGCCTCTCAGTGCACGCGCTCCGGGTGTGGCGCACTGGCCT	1860
OY	1840	-----	1839
Db	1861	TGCAATATGACCCGCTGGAGGATTAATGMAAGGGGAGCTCAGAAAGCTATTATTGGAC	1920
OY	1840	-----TAGGCTGTTCACGTGCAGC	1857
Db	1921	AATGCTGTGTCCTCTTCACATGGTGTGAGCGCGGCTCCGCAATGGCTGTTCATGTCAGC	1980
OY	1858	CTGGCATTAAGCGGTAGTACTCGGATATAGGAGCGCGGCTTGTCAACCGAATGGCGTATC	1917
Db	1981	CTGGCATTAAGCGGTAGTACTCGGATATAGGAGCGCGGCTTGTCAACCGAATGGCGTATC	2040
OY	1918	CCTCCAGGGCAGCTCGGAATGGCGCGTGGCCATCAACGCAAAATCTTGGGCTTCATGCT	1977
Db	2041	CCTCCAGGGCAGCTCGGAATGGCGCGTGGCCATCAACGCAAAATCTTGGGCTTCATGCT	2100
OY	1978	TCTGTGATATTGAGAGCTGCACAAACCTGCATTTCTATTGCTTTGTATACAGTGGCCCAATC	2037
Db	2101	TCTGTGATATTGAGAGCTGCACAAACCTGCATTTCTATTGCTTTGTATACAGTGGCCCAATC	2160
OY	2038	TTGGTGTGAAGCTTAAACATGTTGGGAAACAATTCATCTTAAGGCGTGGGGGTGTGA	2097
Db	2161	TTGGTGTGAAGCTTAAACATGTTGGGAAACAATTCATCTTAAGGCGTGGGGGTGTGA	2220
OY	2098	GGATGGCCAGCTGTGTGCGCTGTGTGGGTGGCGGAACTGTGGGTAGCATTTTAGCTAGCTG	2157
Db	2221	GGATGGCCAGCTGTGTGCGCTGTGTGGGTGGCGGAACTGTGGGTAGCATTTTAGCTAGCTG	2280
OY	2158	GCATATGACAAACGGGGCGCGTGAAGATTGAGCACTTGAGGAACTTATGAAAGCTAGC	2217
Db	2281	GCATATGACAAACGGGGCGCGTGAAGATTGAGCACTTGAGGAACTTATGAAAGCTAGC	2340
OY	2218	GCTTTATATCCACCGTATGCGATTGACGTTGTGTGTAGGCAACCAAGCGGTAGGAAGCGG	2277
Db	2341	GCTTTATATCCACCGTATGCGATTGACGTTGTGTGTAGGCAACCAAGCGGTAGGAAGCGG	2400
OY	2278	AGAGATATGCAATTGCAAAACGCTGTAAAGAAACGGCATATGCTATAGACACTCTGATGTGA	2337
Db	2401	AGAGATATGCAATTGCAAAACGCTGTAAAGAAACGGCATATGCTATAGACACTCTGATGTGA	2460
OY	2338	CCCTTTGGCGAGCCACAGACAGAGAGGTGTGCATACGCGCTTGTAAACAGCATCTTG	2397
Db	2461	CCCTTTGGCGAGCCACAGACAGAGAGGTGTGCATACGCGCTTGTAAACAGCATCTTG	2520
OY	2398	AG 2399	
Db	2521	AG 2522	

RESULT 3
AY055755 2454 bp mRNA linear PLN 31-DEC-2001
LOCUS Chlamydomonas reinhardtii iron-hydrogenase HydA (hyda) mRNA,
DEFINITION complete cds.
ACCESSION AY055755
VERSION AY055755.1 GI:18026269
KEYWORDS
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE
1 (bases 1 to 2454)
Forester, M., Zhang, L., Plummer, S., Ahmann, D., Seibert, M. and
Ghirardi, M. L.
Two putative Fe-only hydrogenases cloned from Chlamydomonas
reinhardtii are coexpressed in cells undergoing anaerobiosis
Unpublished
2 (bases 1 to 2454)
Forester, M., Zhang, L., Plummer, S., Ahmann, D., Seibert, M. and
Ghirardi, M. L.
Direct Submission
Submitted (12-SEP-2001) Basic Sciences, National Renewable Energy
Laboratory, 1617 Cole Boulevard, Golden, CO 80401, USA
Location/Qualifiers
1..2454
/organism="Chlamydomonas reinhardtii"
/strain="21gr"
/db_xref="taxon:3055"
/clone="lam9-1"
1..2454
/gene="hyda"
120..1613
/gene="hyda"
/codon_start=1
/product="iron-hydrogenase HydA"
/protein_id="AAI23572.1"
/db_xref="GI:18026270"
/translation="MSAIVLKCAAVSIRGSSCPARNOVA PAPAIASTRVVALTLEA
PARRIGNVACAAAPAAEPALSHVOQALBELKPKDPTPRKIVCQVAPAVYALLET
LGLPAPATPKQIALEGRLEGFDEVFDTLFGADLTIMESGSLHRLTHLBAHPHSD
BFLPMTSCCPGMIAMLEKSYDPLPYVSSCKSPQMLAMKSYLAEEKGIA PKDMV
MYSIMPCTRKSEADDFECVADPTLRDLQDHYITVEIGLTFKEGIGIALPEDEM
NPMVGSGAGVLFCTTGGMVAALRTAELFTGLPRLPSLSSEVGMGDIETNTIM
VPAQSKFEELKRAABAABAABSTPPPLANDGAGTSEDEGSGITLRVAVANGL
GNAKULITMGAQEAHYDFVEIMACPAAGCGGQPRSTDKA1TQRQALALYNLDEKS
TLRSHENSIRELYDTYLGEPVGHAKHLHTHYVAVGVEBERDEKK"

BASE COUNT 493 a 747 c 784 g 430 t
ORIGIN

Query Match 88.3%; Score 2119.4; DB 8; Length 2454;
Best Local Similarity 93.7%; Pred. No. 9, 2e-253;
Matches 2280; Conservative 0; Mismatches 6; Indels 147; Gaps 1;

QY 40 CTCGAACCCCTCGAAGCTTTTCCACAGTTTACACCCCAATTCGAGCGCGCTCCAAAG 99
DB 1 CTCGAACCCCTCGAAGCTTTTCCACAGTTTACACCCCAATTCGAGCGCGCTCCAAAG 60
QY 100 TCGCTCGTTGCTCTTCATCGACCACTATTATTCTAATATCTGAAGACGGACAGA 159
DB 61 TCGCTCGTTGCTCTTCATCGACCACTATTATTCTAATATCTGAAGACGGACAGA 120
QY 160 TGTGCGCGCTCGTCTGAAGCCCTGCGCGCGCTGTCTATTCGCGGCAAGTCTCTGAAGG 219
DB 121 TGTGCGCGCTCGTCTGAAGCCCTGCGCGCGCTGTCTATTCGCGGCAAGTCTCTGAAGG 180
QY 220 CGGAGAGGTGCGCGCGCGCTCGCTCGACGACGACCGGTGATGAGCCCTTGCA 279
DB 181 CGGAGAGGTGCGCGCGCGCTCGCTCGACGACGACCGGTGATGAGCCCTTGCA 240
QY 280 CACTGAGGCGCGCGCGCTGAGCAAGTCTGCGCGCGCTGCGGCAAGCCGCTG 319

DB 241 CACTGAGGCGCGCGCGCTGAGCAAGTCTGCGCGCGCTGCGGCAAGCCGCTG 300
QY 340 CGGAGAGGTGCGCGCGCTGAGTATCCAGAGGCGCTGCGGACCTTCCAAAGCAAGAG 399
DB 301 CGGAGAGGTGCGCGCGCTGAGTATCCAGAGGCGCTGCGGACCTTCCAAAGCAAGAG 360
QY 400 ACCCGACGCGCAAGAGCTCTGCGTGAAGTGGCTCCGCGCTGCTGCTGCTATTCGCG 459
DB 361 ACCCGACGCGCAAGAGCTCTGCGTGAAGTGGCTCCGCGCTGCTGCTGCTATTCGCG 420
QY 460 AGACCTTGCGCTGCG 519
DB 421 AGACCTTGCGCTGCG 480
QY 520 GCCTGCGCTTGAACAGAGTGTGACAGCTGTTGGGCGCGACCTGACCATCATGAGAG 579
DB 481 GCCTGCGCTTGAACAGAGTGTGACAGCTGTTGGGCGCGACCTGACCATCATGAGAG 540
QY 580 AGGAGAGGAGCTGCTGCAACCGCTCACCGAGCACTGGAAGCCCAAGCGCACTCGAAG 639
DB 541 AGGAGAGGAGCTGCTGCAACCGCTCACCGAGCACTGGAAGCCCAAGCGCACTCGAAG 600
QY 640 AGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 699
DB 601 AGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 700 ACCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759
DB 661 ACCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 760 TGTGCAAGTCTTACCTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 819
DB 721 TGTGCAAGTCTTACCTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 820 CCATATGCTGCTGCAACCGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 879
DB 781 CCATATGCTGCTGCAACCGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 840
QY 880 CCGACCCCAACCTGCGCGCGAGCTGCAACGCTGATCACCACCGTGGAGTGGGCAATCT 939
DB 841 CCGACCCCAACCTGCGCGCGAGCTGCAACGCTGATCACCACCGTGGAGTGGGCAATCT 900
QY 940 TCAGAGAGCGCGGATCAACCTGCGCGAGCTGCGCGAGGCGAGTGGGCAATTCGAATG 999
DB 901 TCAGAGAGCGCGGATCAACCTGCGCGAGCTGCGCGAGGCGAGTGGGCAATTCGAATG 960
QY 1000 GCGTGGGCTGCGGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1059
DB 961 GCGTGGGCTGCGGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1060 TGGGACGCGCTATAGCTGTTACAGGCGCGCGCGCTGCGCGCGCTGAGCGGAGG 1119
DB 1021 TGGGACGCGCTATAGCTGTTACAGGCGCGCGCGCTGCGCGCGCTGAGCGGAGG 1080
QY 1120 TGGGCGGATGAGAGGAGATCAAGAGAGCAACATCATCATGATGCTGCGCGCGCGG 1179
DB 1081 TGGGCGGATGAGAGGAGATCAAGAGAGCAACATCATCATGATGCTGCGCGCGCGG 1140
QY 1180 AGTTGAGAGGCTGCTGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1239
DB 1141 AGTTGAGAGGCTGCTGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
QY 1240 CCGCGCGCGCGCTGCG 1299
DB 1201 CCGCGCGCGCGCTGCG 1260
QY 1300 GCATCACTGCGCGCTGCG 1359
DB 1261 GCATCACTGCGCGCTGCG 1320
QY 1360 AGATGAGGCGCGCGAGGCAAGTGAATGCTTGTGAGATCATGCTGCGCGCGCGCT 1419

Db	556	GACCATCATGGAGAGAGGACCGGAACTGCTGCACCGCTTCAAGAGACACTTGAAGGCCA	615
OY	626	CCCGCACTCCGACGAGCCGCTGCCCCATGTTCAACCAAGCTGCTGCCCGGCTGATCGTAT	685
Db	616	CCCGCACTCCGATGAGCGGCTGCCATGTTTCAACAGAGCTGTGCCCGGGTGGTGCCAT	675
OY	686	GCTGGAGAAATCTTACCCGAGACTGATCCCTTAAGTGAAGAGCTGCMAAGCCCCAGAT	745
Db	676	GATGAGAGAACTCTACCCCTGAACTCATCCCTTTCGTAGCTCTCGAAGAGTCCCAAT	735
OY	746	GATCGTGGGCGCATGATGATCAATCTTACCTTAGCGGAAAAGAGGGCATGCGCCAAAGGA	805
Db	736	GATGATGGGCGCCATGCTGAAGAACCTTACTGTCCGAAAGCAGGAGCATCCCGCCAAAGGA	795
OY	806	CATGATCATGCTGTCTCATCTATGCTCCCTGCACGCGCAAGCAGTCCGAGGCTGACCGGACTG	865
Db	796	CATGTCATAGCTCTCGTATATGCTTTGGGTGGCGAAGCAGGGGCGAGGCTGACCGCGAGTG	855
OY	866	GTTCTGTGTGACGCCGACCCCACTTCGCGCACTGACCAACGTCATATCAACAGCTGGA	925
Db	856	GTTCTGTGTG--AAGCAGACCCGCGGTGGCGACAGTGAACACGTAATACCAACCGCCGA	912
OY	926	GCTGGGCAACATCTTCAAGGAGCGGGGACATCAACTGGCGGAGCTGCCGAGGGGAGTG	985
Db	913	GCTGGGCAACATCTTCAAGGAGCGGTGCAATCACTGCGCGAGCTGCCGCAAGTACTG	972
OY	986	GGACAAATCCAAATGGGCTGTGGGCTCCGGGCGCGAGCTGCTGTTCCGACACCGCGGTGT	1045
Db	973	GGACCAAGCGCTGGGCGCTGGGCTCCGGCGCGCGGCTACTGTTCCGACACCGCGGCGT	1033
OY	1046	CATGAGGCGGCGCTGCGCACCGGCTTATGCTGTTACGCGGACAGCGCTGCGCGCTT	1108
Db	1033	CATGAGAGCGGCGCTGCGCACCGGCTTACGAGATAGTGAACAGAGACCGCTCGCGCTT	1092
OY	1106	GAGCTTGAAGCAGAGTGGCGCGGATGGAAGCGCATCAAGAGACCAACATATCAATGTGTCC	1162
Db	1093	CAACTTGAAGCGAGTGGCGGCTTTGAGACGGCATCAAGAGAGGCGTCCGTGACCTGATGCC	1152
OY	1166	CGCGCCCGGGTCCAGATTGAGAGACTGTGTAAGCACCGC-----	1202
Db	1153	CGCTCCGGGCTTCAAGTTGCGCGAAGTGTGGCGGAGCGCTGGCGCACAAAGTGAAGGA	1212
OY	1206	---GCGCGCGCGCGCGCGAGGCGCGCGCGACCGCACCCCCGGGCGC---TGCGCTG	1258
Db	1213	GCGCGCGCGGCTGAAGCGCGCGCGCGGTGAAGGGCGCGTGAAGCGCCATCGCGTA	1272
OY	1259	GGAACGGCGCGCGGGCTTACCAAGCGAGGACGGCAAGGGCGGCATTCACCTGCGCTGGC	1318
Db	1273	CGACCGCGCGCGAGGGTCTTCTCCACGGAATGACGGCAAGGCGCGCTGAAGCTCGAGTGGC	1332
OY	1319	CGTGGCCAAACGGGCGTGGGCAAGGCCAAGAGTGAATCAACAGATGACAGCGCGGAGGCG	1372
Db	1333	GGTGGCGAAACGCGCTGGGCAAGGCCAAGAGCTGAATGGCAAGATGATATCTGGCGAGGCG	1392
OY	1379	CAAGTACGACTTTGTGAGATCATGGCTGGCCCGCGGCTGTGTGGGCGCGCGGCCA	1433
Db	1393	CAAGTACGACTTCTGTAAGAAATATATGGCTGGCCCTGCGGCTGCTGGGCGGCGGCCA	1452
OY	1439	GCCCGGCTTCACGACAGAGCCATTAACGCAAGAGCGGCGCGGCTGTATACACTTGA	1498
Db	1453	GCCCGGCTTCACGACAGAGCATCAACCAAGAGCGGCGCGGCTGTATACACTTGA	1512
OY	1499	CGAGAGTCTCAGCGTGGCGCGGACGACAGAAACCCGTTCATCCGCGAGCTGTACAGAC	1558
Db	1513	CGAGGCAACAGCTGCGCGGAGGCCACAAAACAGAGGGGTTCACCAACTGTATCAAGGA	1572
OY	1559	GTACCTCGGAGAGCGGCTGGGCGACAAAGGCGACGAGCTGTGCACCACTACGTGGC	1612
Db	1573	GTTCCTGGGCGAGGCCCTGTCCACCGCGCCACAGAGCTGTGCACCACTACGTGGC	1632
OY	1619	CGCGCGCTGAGG 1632	

DB 1633 AGCGCGCCGAGG 1646

RESULT 5
CFU298228 2420 bp mRNA linear PLN 04-JAN-2002

LOCUS AJ298228
DEFINITION Chlorella fusca mRNA for Fe-hydrogenase (hya gene).
ACCESSION AJ298228
VERSION AJ298228.1 GI:18073434
KEYWORDS Fe-hydrogenase; hya gene.
SOURCE 'Chlorella' fusca.
ORGANISM Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Chlorococcales; Scenedesmeceae; Scenedesmus.

REFERENCE 1
AUTHORS Winkler, M.
TITLE The Fe-hydrogenase of Chlorella fusca
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2420)
AUTHORS Happe, T.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-2000) Happe T., Molekulare Biochemie, Botanisches Institut der Universitaet Bonn, Karlrobert-Kreitzen-Str.13, Bonn, NRW, 53115, GERMANY

FEATURES
source
1..2420
/organism="Chlorella' fusca"
/db_xref="taxon:3073"
67..1377
/gene="hya"
67..1377
/gene="hya"
/EC_number="1.18..99.1"
/function="hydrogen metabolism"
/codon_start=1
/product="Fe-hydrogenase"
/protein_id="CAC83291.1"
/db_xref="GI:18073435"
/translation="MCCPVVASRHAGRAHVAVRAAGPTSECDCEPTPOAKLPHMOQ
LDELAKRESRLMLAIOIASAVRAVAIAETGLAGDVTIGLVYGLMGLPDYEDTIT
FGALDTMBESTELHRLQHLBHPHKEIPDPFTSCCGRVAMVEKSNELPYLS
SCSPQMLGAVINRYAQQGVQPSIDCVASVMPCYRKQGEADREWPNFTGGLARI
VDHYVTTAEVGIKLELPESNFNPIDEGGALPLTTGGVMAALRTVY
EVVYQKPMGRVDEEVGLDGLKEAEITLKPQDPSPKAFAGADGGITLKIANGI
GNAKKLILSLEGGAKYDPIEVNACPGCGCGGCGPSRTPDKOILQKQQAAMYNLDEERS
TIRSHNPRTGALYDKFLGAPNSHKADLHTHYVAGIPEEK"

transit_peptide 67..129
/gene="hya"
130..1374
/gene="hya"
/product="Fe-hydrogenase"
/EC_number="1.18..99.1"
/mat_peptide
mat_peptide 130..1374
/gene="hya"
/product="Fe-hydrogenase"
/EC_number="1.18..99.1"

BASE COUNT 578 a 581 c 665 g 592 t 4 others

ORIGIN

Query Match 18.7%; Score 448.6; DB 8; Length 2420;
Best local Similarity 61.7%; Pred. No. 1.6e-46;
Matches 807; Conservative 0; Mismatches 409; Indels 93; Gaps 2;

331 CACCGCTGGGAGGCGCTTTGATGATCTCCAGACGCGCTGCCGAGCTTCCCAAGC 390
DB CTCGACACCTCAAGCCCAAGCTGCTCACTGCGACGAGCGCTTGATGAGCTCGCAAGC 217
391 CCAAGGACGACCCCAAGCGCAAGCAGCTCTCCGTGACAGGTGCTCCGCCCTTGTGTGTC 450
DB CCAAGGACGACCCCAAGCGCAAGCAGCTCTCCGTGACAGGTGCTCCGCCCTTGTGTGTC 271
451 CTATTGCGGACGACCTGGGCGCTGGCGCGGCGCGCACACCCCAAGGAGCTGGCGGAG 510
DB CTATTGCTGAACCATTTGGCTTGGCCCCAGAGATGTCACCATTTGGGCGAGCTTGACTG 331
511 GCCTCCGCGCGCTGGCTTTGACGAGGTGTTGACAGCGCTGTTGGCGCCGACCTGACCA 570

332 GGCTGCGTATGCTTGGCTTTGATTGATGCTTTTGAGACCCGTTTGGTGCTGACCTGACA 332

333 TCAATGAGAGAGGAGACAGCAAGCTGCTGCAACCGCTTCAACGAGCACTGAGAGCCOACCCG 630

334 TATATGAGAGAGGAGAACGAGAGCTCTGCTCATGCGCTTGACAGACCAATCTGGACAGACCCCA 451

335 ACTCCAGCAGAGCCGCTGCGCCATTTACACAGCTGCTCCCGGCTGGATGCTATGCTG 690

336 ACAAGAGAGAGGCCACTGGCCCATTTACCAAGTTGCTGCCAGGCTGGGTTGCCATGGTTG 511

337 ACGAAATCTTACCCGGAGCTGATCTCCCTTACGTGAGCACTGCAGAGAGCCCAAGTATGTC 756

338 AAAAGAGCAATCTTGAGCTCATCTCCCTTACTGTCATCTTGCAGTGCCTCAATGATGTC 571

339 TGGCGGCAATGCTCAAGTCCCTACTAGCGGAAAGAGGAGCTGCGGCAAGAGCACTG 810

340 TTGGGGCCGTTATCAAGAACTACTATGACAGCAGGTTGAGTGCAGCCCAAGTACATCT 631

341 TCATGGTGTCCATCATGCTCCCTGACGCGCAAGAGTGCAGAGGCTGACCCGACTGTTCT 870

342 GCAACCTGTCTAGTCATGSCATGCGTACGCAAGCAGGAGAGGCTGACCCGAGTGGTTCA 691

343 GTGTGAGACCCGAGCCCAACCTCTGCGCAGCTGAGACCACTGATCAACCACTGTGAGCTG 930

344 ACACCAAGGTGAGGCTTTGGCCCGTATGTTATGATGTGACTCATCTGCTGAGGTTG 751

345 GCAACATCTTCAAGAGCGCGGATCAACTCTGCGGCTGACCCGAGGCGAGTGGAGCA 990

346 GTAAATATTTCTGGAGCGTGGCATCAGTGAATGACTGCTGCCAGAGCAACTTTTGACA 811

347 ATCCATATGGGCGGTGGCTCGGCGCGCGGCGCTGTTCCGACACACCGCGGTCTATGG 105

348 ACCCCATTGAGGAGGAGCAAGGTGTGCTCTGCTGTTTGGCACACTGGAGGTGATGG 871

349 AGGCGCGCTGCGCAAGCGCTATGAGCTGTTCAAGGAGCAGCGGCTGCGCGCTGAGCC 111

350 AAGCAGCACTTGGCAAGCTTATGAGTGTGATGAGTCCAGAAAGCCATGGTCTGCTTGACT 931

351 TGAAGAGTGGCGGCGCATGAGCAGCATCAAGAGACCAATACATGTTGCCCGC 117

352 TTGAGAGGTGGGAGGCTTTGAAGGATCAAGAGGACAGATCACTCAACCCAGAG 991

353 CCGGATCMAAGTTTGAAGAGCTGTGAGAGCAGCGCGCGCGCGCGAGGCGCG 123

354 ACGACAGCCCATTTCAAGCTT----- 101

355 CGCACGCGCACCCCGGGGCGCTGGCCTGGAGCGGCGCGCGGCTTACCAAGCAGAGAG 129

356 -----CGCAGAGCTG 102

357 GCAGGGGAGGCATCAACTGCGGTGGCCGTGGCCCAAGGGCTGGGCAACGCCAAGAAC 135

358 ATGGGAGGAGCATCAAGCTTCAAGATTGACGTAGCCCAATGGAGCTTGGCAATGCCAAGAAC 108

359 TGATCAACCAAGATGCAAGCGCGGAGGCCAAGTACACTTGTGAGATCATGCGCTGCG 141

360 TCATCAAGAGCTGTGAGAGGGCAAGCCAAATGATTTCAATTGAGGTCTATGGCATGGC 114

361 CCGCGGGCTGTGTGGGCGGCGGCGCAAGCCGCTTCCACCGAAGGCCATACGCAAG 147

362 CTGTGTGGCTGCAATGGCGGAGCGGTGCAACCCCGCAGTACTGCAAGAGATCTTCGACA 120

363 ACGGAGGCGGCGCTGTGCAACCTGAGAGAGAAATCCAGCTGCGCGCGAGCCAGAGAG 153

364 ACGCCAGCAGAGCTATGTACAACTGATGAGCGCAGTACCATCCGCGGAGCCATGATTA 126

365 ACCCATTCATCAAGGCGCTGTATGACAAATTTCTAGGCGCACCCCAACAGCACAAGGAC 132

366 ACGAGCTGTGACACCCCACTACGTGGCCGCGCGCGCTGAGAGAGAAAGGA 1639

367 ATGATCTGTGACACACACTATGTGGCAGGTGGAATTTCCAGAGGGGAA 1373

RESULT 6	LOCUS	1607 bp	MRNA	linear	PLN 15-MAR-2001
AF276706	Scenedesmus obliquus Fe-hydrogenase mRNA, partial cds.				
DEFINITION	Scenedesmus obliquus Fe-hydrogenase mRNA, partial cds.				
ACCESSION	AF276706				
VERSION	AF276706.1				
KEYWORDS	GI:12581497				
SOURCE	Scenedesmus obliquus.				
ORGANISM	Scenedesmus obliquus				
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Chlorococcales; Scenedesmaceae; Scenedesmus.				
AUTHORS	1 (bases 1 to 1607)				
TITLE	Munichler, R., Stangier, K., Senger, H. and Schulz, R.				
JOURNAL	Molecular evidence for a Fe-hydrogenase in the green alga Scenedesmus obliquus				
MEDLINE	Curr. Microbiol. 42 (5), 353-360 (2001)				
PUBMED	21292711				
REFERENCE	11400057				
AUTHORS	2 (bases 1 to 1607)				
TITLE	Munichler, R., Stangier, K., Senger, H. and Schulz, R.				
JOURNAL	Direct Submision				
FEATURES	Submitted (09-JUN-2000) Botanisches Institut, Christian-Albrechts-Universitaet, Olshausenstr. 40, Kiel 24098, Germany				
source	Location/Qualifiers				
CDS	1..1607				
	/organism="Scenedesmus obliquus"				
	/strain="wild type D3"				
	/db_xref="taxon:3088"				
	<1..1212				
	/codon_start=1				
	/evidence="not experimental"				
	/product="Fe-hydrogenase"				
	/protein_id="AAC59621.1"				
	/db_xref="GI:12581498"				
	/translation="PRHQQTLDLAKRKWKMTAOTAPVRGIAETMGINPDTYVGQMTVGLGPDVFDTLFGADTLTDEGLTDLHRLODHLDEHPNKEPILPMTSCCPGWAVMKSINPELPIYLSCKSPQPMWGLAVIKYFAEAQAKPDICNVEMPTVRSGLEAPRSGSTHHRAGRDVDVMTTALGKIPVERGKINLELSPDNVGEISGGGGLPFTTGTVMEALRTVEYTAALGPSSLTSTAMPQAGASRPSPOAPTPASRPLOAOTESGITLINIAVANGGNKKIKIOLLAGESYDPTVWACGGGIGGGGPORRNIOLOKROAAWYDIDERAVIRTRTEPLGALYKELGRPNHKAHLEHTHYVAGGVPRRSCEAM"				
	292..375				
	/note="Region: H-cluster 1"				
	451..486				
	/note="Region: H-cluster 2"				
	952..999				
	/note="Region: H-cluster 3"				
BASE COUNT	330 a 450 c 520 g 307 t				
ORIGIN					
Query Match	17.7%; Score 424.6; DB 8; Length 1607;				
Best Local Similarity	64.2%; Pred. No. 1.6e-43;				
Matches 816; Conservative	0; Mismatches 359; Indels 96; Gaps 8;				
363	CAGACGGCGCTCGCCGACGCTTGGCCAGGCCAGGACACCCACGCGCAACGACGCTTGC	422			
10	CAGACGACGCTGAGTAGCTAGCCACACCTTAAGA-----GAGGAAGTGATGATC	60			
423	GTGCAGATGCTCGCGCGCTTCGTGTGCTATTGGCCGAGACCTGGGCGCTGGCGCGGC	482			
61	GCCCAAGTTCGACGACGAGTGCCTGGC--ATTGCAAGACCATGGGCTCAATCTCGGG	117			
483	GCCACCAACCCCAAGACGCTGGCCGAGGGCTCGCGCTCGGCTTTGACGAGTGT	542			
118	GATGTGACAGTGGCCAGATGATGACCGGCTGGCATGTGGGCTTTGATTAATGTGTT	177			
543	GACACGCTGTTGGCGCCGACCTGACCATGAGAGAGGGCAGCGACTGCTGCACCGC	602			
178	GACACGCTGTTGGTGTGACCTTCAACATATGAGAGAGGGCACAGAGCTGTCACAGG	237			

OY	603	TTACACGAGCAKCTTGGAGAGCCCAACCCGGACTCCGAGAGAGCGCTCCCATGTTACACAGC	662
Db	238	CTTCAGAGACACACTGGAGACAGCAACCCCAACAAGAGAGAGCCCTCCCATGTTACACAGC	297
OY	663	TGCTGCCCCGCTGGATCGCTATGCTGGAGAAATCTTACCCGGACTGATCCCTTACGTC	722
Db	298	TGCTGCCCCGCTGGATCGCATGCTGGAGAAATCTTACCCGGACTGATCCCTTACGTC	357
OY	723	AGCAGCTCAAGAGCCCCAGATGATGCTGGAGGCGCATGTCAGTCCCTCACTAGCCGAA	782
Db	358	TCTTCTGCAAGTCAACCCAGATGATGCTGGAGGCGCATGTCAGTCCCTCACTAGCCGAA	417
OY	783	AAGAGGGGCATTCGCGCCAAAGGACATGCTCATGCTGTCATATGACCCCTGACGCGCAAG	842
Db	418	GAGGCGGCGCCAAACCCGAGGACATGCTATCATGTAAGCTGATGCTCCCTGGGTGGCAAG	477
OY	843	CAGTCGGAGGCTGACCGGCACTGTGTTCTGTGTGAGCGCCGACCCCACTCTGGCCACTG	902
Db	478	AGCGGCGAGGCGGAAACCGCGAATGTTATCAACACCAAGGGCTGAGCGACGCGCACTG	537
OY	903	GACCACTGATCATCCACCGTGGAGCTGGGCAACATCTTCAAGAGAGCGGACATCAACTG	962
Db	538	GACCATGTTATGACACACGCGAGCTGGCAAGATCTTTGAGAGCGGGAATCAAGCTG	597
OY	963	GCCGAGCTGCCCCGAGGCGGAGTGGGACATCCATGGGCGTGGGCTCGGAGCGCGGCGTG	1022
Db	598	AACGAGCTGACAGAGTGGCGTTTGACAAACCCGTAAGCGAGGCGAGCGGCGGCGCTG	657
OY	1023	CTGTTTCGGGACCAACCGGCGGTGTATGAGAGCGGCGGCTGCGGCAACGCTTATGAGCTGTT	1088
Db	658	CTGTTTGGGACCACTGGGCGGCGTGTATGAGAGGCGGCGTGGCGACGTTTACGAATGCTG	717
OY	1083	ACGGGCAAGCGGCTGCGGCGCTGAGGCTGAGCGTGAAGTGGCGGATGAGACGGCATCAAG	1142
Db	718	ACAGGAGAAAGCCCT-----TGGACCGAGGATGCTTTGACGACATCCAC	764
OY	1143	GAGACCAACATCACATGATGTCGCCGCGCCGGGTCGAAGTTGAGAGACTGTTAGACAC	1202
Db	765	GGC-----CTGGACCCG	776
OY	1203	CGGCGCGCGCGCGCGGCGGAGCGCGCGGCGACACCCCGGCGCGCTGGGAC	1262
Db	777	AGCACAAGGCGCTTGGCCACGACCTTACCCAGGCGCCCAACAGCC-----	823
OY	1263	GCGCGCGCGGCTTGCACGACGCGAGGACGCGAGGCGGCGATCACACTGGCGCTGGCGTG	1322
Db	824	--CTTCAAGGCTTTGACAGGCGGCAACGG--AGTCCGGATATACCTCAAACTTGTGTC	879
OY	1323	GCCAAACGGGCTTGGGCAACGCCAAGAAGTGTATCCAAAGATTCAGAGCGGCGGCGCAAG	1388
Db	880	GCCAAACGGGCTTGGGCAACGCCAAGAAGTGTATCCAAAGCTGAGCTGACAGGCGAGGCAAG	939
OY	1383	TACGACTTTGTGAGATCATGGGCTGGCGCGGCGGCGCTGTGTGGGCGGCGGCGGCAACCC	1442
Db	940	TACGACTTTCACGAGGTCAATGGCTTGGCGCGGCGGCGCTGTGTGGGCGGCGGCGGCAACCC	999
OY	1443	CGCTCCACCGCACAGGCGCATCAACGAGAAGCGGACGCGCGCTGTACACTGTGACACAGTAC	1502
Db	1000	CAACGGA---ACAAACAGATCTCTGACAGAAGCGCCACAGCGGCGCATGTATGACTGACG--	105
OY	1503	AAGTCAACGCTGCGCGCGACGACGAGAACCCGCTGATTCGCGAAGCTGTACACACAGTAC	1566
Db	1055	-AACGCGCGGTATCCGCGGCAACGAGAAACCGCTGATTTGGTGGCTGTATAGAGATTG	1113
OY	1563	CTCGGAGAGCGCGCTGGGCGCAAGGCGGACGAGCTGCTGCAACCCATCACTGGCGGCG	1622
Db	1114	CTGGGAGAGCCCAACGCGCAACAGGCGGACGAGGCTGTCTGCAACGCAATTACCTGGCGGCG	1172
OY	1623	GCGCTTGGAGGA	1633
Db	1174	GCGGTGCGCGA	1184

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
DEU07229	Desulfovibrio fructosovorans potential NAD-reducing hydrogenase	U07229	U07229.1	GI:466362		Desulfovibrio fructosovorans.	Desulfovibrio fructosovorans Bacteria; Proteobacteria; delta subdivision; Desulfovibrionaceae; Desulfovibrio.	Malik, S., Salmamime, I., De Luca, G., Rousset, M., Dermoun, Z. and Belaich, J. P.	Characterization of an operon encoding an NADP-reducing hydrogenase in Desulfovibrio fructosovorans	J. Bacteriol. 177 (10), 2628-2636 (1995)	95270577	7751270	2 (bases 1 to 5251)	Malik, S.	Direct Submission
															Submitted (10-MAR-1994) Souria Malik, Bioenergetique, Ingenierie des Proteines, Centre National de la Recherche Scientifique, 31 chemin Joseph Aiguier, Marseille 13402, Cedex 20, France
															Location/Qualifiers
															1..5251
															/organism="Desulfovibrio fructosovorans"
															/db_xref="taxon:878"
															85..90
															109..114
															369..376
															383..398
															/note="potential NAD-reducing hydrogenase subunit"
															/codon_start=1
															/transl_table=1
															/protein_id="AA07054.1"
															/db_xref="GI:466363"
															/translation="MONSTCOANGECVPEHVAUPLQVLYEVQFISLPQKBEHLTVLHAGVFGFLPLEVOQVADHMEVPLAQVGVASTYTFPTWPKGKTPISVCHGTHCFVAGADKVVAAHFEQDKIDIGTVPGRFSIDLRCVGGCALAPIVMGKVGYNVPCQVKKILAEY"
															925..929
															937..1317
															/note="potential NAD-reducing hydrogenase subunit"
															/codon_start=1
															/transl_table=1
															/protein_id="AA07056.1"
															/db_xref="GI:466365"
															/translation="MAATTEKKQRIATRNCGFIDPESIDVIALRGYBGLAKVLTTPRAVLDVKSGRGGAAGFPPTGIMKGLAGKAOKVWCADEGDFEPMRAVILEGDSVSEAMALGIVATIGATRGVVYIARPELAIKILKKAIDAREVGLIGENIFSGEGDDILKTCAGAFVCGEETALIRSNBGRGGEVYTPPPAOSGYEKPTVANNVETFANIPIIINGADWFSIGTATSKGTVPFALAGKIQVGLIEVWGSLREVIIDIGCGCPGKAFAVOTGSPSGCALANDLVAIDYELAACSKIMSGVWMDDECOMVSAKPLFTMETGCTGCTPCRIKSLYEILDRITKGTGADIDRLKSLSEIHKDTPTLCGGMMPNLIISTMDTPANEYEAHVADKCPAHVCTALITLITPIDPAKTCGGCTCTCVPGECSGTGKKQHTIDTTRCICKCACYDKCKRDSIIKKQ"
															2823..2827
															2834..4551
															/note="potential NAD-reducing hydrogenase subunit"


```

exon /number=3
      /260..2325
      /gene="hyda"
      /number=4
      /2326..2621
      /gene="hyda"
      /number=4
      /2622..3089
      /gene="hyda"
      /number=5
      /3090..3400
      /gene="hyda"
      /number=5
      /3401..3877
      /gene="hyda"
      /number=6

BASE COUNT 1002 a 1335 c 1474 g 1190 t
ORIGIN
Query Match 10.2%; Score 244.6; DB 8; Length 5001;
Best Local Similarity 70.1%; Pred. No. 2,1e-21;
Matches 344; Conservative 0; Mismatches 144; Indels 3; Gaps 1;

QY 599 CCGCTCACCGAGACCTGGAGGCCACCCGCACTCCGACGAGCCGCTGCCATGTTAC 658
DB 2612 CTGCTTGACGAGACACACCTGGAGGACCCCAAGAGAGAGCCCTGCGCATGTTAC 2671
QY 659 CAGCTGCTGCCCCGGGTGATGCTATGCTGAGAAATCTTACCGGACCTGATCCCTTA 718
DB 2672 CAGCTGCTGCCCCGGGTGATGCTATGCTGAGAAATCTTACCGGACCTGATCCCTTA 2731
QY 719 CGTGAACAGCTGCAAGAGCCCGACATGATGCTGCGCCATGATCAAGTCTTAAGCT 778
DB 2732 CGTGTCTTCTGCAATGCGCCCAATGATGCTGCGCCATGATCAAGTCTTAAGCT 2791
QY 779 GGAAGAAAGAGGATGCGCCCAAGAGCATGATGCTGCTCATATGCTGCTGACGGC 838
DB 2792 TGCCTGAGGCGGCGGCGGAGCTGAGGACATCTGCAAGCTGAGCGTATGCTGCTG 2851
QY 839 CAAGCAGTGGAGGAGCTGACCGGCGACCTGCTGCTGAGACCGGACCCCACTGCGCA 898
DB 2852 CAAGCAGGCGGAGGAGCTGACCGGCGACCTGCTGCTGAGACCGGACCCCACTGCGCA 2908
QY 899 GCTGACACAGCTATCACCACCGTGGAGCTGGGCAATCTTCAAGAGCGCGCATCA 958
DB 2909 CGTGAACAGCTATCACCACCGTGGAGCTGGGCAATCTTCAAGAGCGCGCATCA 2968
QY 959 CCTGCGCAGCTGCGCGGAGGCGAGTGGGCAATCCATGAGGCGTGGCGCGCGCG 1018
DB 2969 GCTGAACAGCTATCACCACCGTGGAGCTGGGCAATCCATGAGGCGTGGCGCGCGCG 3028
QY 1019 CGTGTCTTCTGCAATGCGCCCAATGATGCTGCGCCATGATCAAGTCTTAAGCT 1078
DB 3029 CGTGTCTTCTGCAATGCGCCCAATGATGCTGCGCCATGATCAAGTCTTAAGCT 3088
QY 1079 GTTACGCGCA 1089
DB 3089 GTTACGCGCA 3099

RESULT 9
LOCUS DFHYDABG 2887 bp DNA linear BCT 30-MAR-1997
DEFINITION D. fructosovorans hyda and hydB genes.
ACCESSION Y11759
VERSION Y11759.1 GI:1914863
KEYWORDS Fe-hydrogenase; hyda gene; hydB gene; iron-only hydrogenase.
SOURCE Desulfovibrio fructosovorans.
ORGANISM Desulfovibrio fructosovorans
Bacteria; Proteobacteria; delta subdivision; Desulfovibrionaceae;
Desulfovibrio.
REFERENCE 1 (bases 1 to 2887)
AUTHORS Casalo, L., Hatchikian, E. C., Forget, N., de Philip, P., Belach, J. P.

```

```

TITLE and Rousset, M.
JOURNAL Molecular study and partial purification on iron-only hydrogenase
AUTHORS in Desulfovibrio fructosovorans
TITLE Unpublished
JOURNAL 2 (bases 1 to 2887)
AUTHORS Direct Submission
TITLE Submitted (11-MAR-1997) L. Casalo, CNRS, UPR9036 BIP03, 31, Chemin
JOURNAL Joseph Aiguier, Marseille, 13402 Cedex 09, FRANCE
FEATURES
source location/Qualifiers
1..2887
/organism="Desulfovibrio fructosovorans"
/strain="DSM 3604"
/db_xref="taxon:878"
445..1721
/gene="hyda"
445..449
/gene="hyda"
456..1721
/gene="hyda"
/EC_number="1.12.2.1"
/function="Fe-hydrogenase large subunit"
/codon_start=1
/transl_table=11
/product="cytochrome-c3 hydrogenase"
/protein_id="CAA72423.1"
/db_xref="GI:1914864"
/db_xref="SPTREMBL:O08311"
/translation="MSRIEMAKIFPEQIVYPPPGTNDQAYIVQVDETKICGDTCHGY
CPTGATIGSGEPKRVVDPACINCGCLTHCPVAIVETVSFVEIEKLIKDKNVK
IAMPAPVAVLADPFGMPGLGAVTTEHMLTGKQGFDMVWNEFTADVTIIEGSEL
LARITKLDKPLPOFTSCCPGOKYAEIYFELPHFSSCKPIGMMGLATYGAKE
LGYEKROIYVSIIMPCTAKKPEKGRMENDASFRDIDATINRELAVYMKKGIIDPK
IANGKRDVMEGSTGATIFGYSGVMEALAPFAVQALTKRPSMDPFAVAGLNIK
EATINIGTIDVAVVAVNGKAPKACDEYKAKGSPYHFIEMACREGCTMGSGQPIMP
TVLESNRRITTKFYASLKRRLVDAQA"
1814..2198
/gene="hydB"
1814..1819
/gene="hydB"
1824..2198
/gene="hydB"
/EC_number="1.12.2.1"
/function="Fe-hydrogenase small subunit"
/codon_start=1
/transl_table=11
/product="cytochrome-c3 hydrogenase"
/protein_id="CAA72424.1"
/db_xref="GI:1914865"
/db_xref="SPTREMBL:O08312"
/translation="MSIILATTRRGFMKTRACVLTGGALIGRLTSKAVAAKOLKXYM
DRVNAVYGAADAKFKYRASODNAOVITLVKKFLHEPLSHESEHLITTKVDRSKDLAAL
KAKGLYPTPRAKEFEKPYPO"
1824..1925
/gene="hydB"

BASE COUNT 613 a 985 c 776 g 513 t
ORIGIN
Query Match 10.1%; Score 242.4; DB 1; Length 2887;
Best Local Similarity 55.4%; Pred. No. 4,5e-21;
Matches 555; Conservative 0; Mismatches 411; Indels 36; Gaps 3;

QY 368 GCGGCTCGCGGAGCTTGCCAGCCAGAGACACCCAGCGCCAGCAGCTGCGCTGCA 427
DB 710 GTCTCTGTCGCCGAGATGAGCCAGCTCAAGACAAAGAACTCAAGCTATCCCAT 769
QY 428 GGTGCTCGCGGCGTTCGTGCTATGCTGAGACCTTGCGCGCGCGCCAC 487
DB 770 GCGGCGCGCGCGCTGCTGCTAGCCCTGCGCATCTTCCGATGCCCTCGCGCGG 829
QY 488 CACCCCAAGCAGCTGCGCGGAGGCGCTCGCGCTGCGCTTGAAGAGGTGTTGAC 547
DB 830 CACCAACGACATGCTCAGCGGCGCTCAAGCAGCTTGCTGCAACAGTCTGAGCAA 889

```

```

QY 548 GCTGTTTGGCCGCGACCTGACCATCATGAGAGGCGACGAGCTGCTGACCGGCTCAC 607.
DB 890 CGAGTTACAGCGCGCAGCTCACCATCTGGAGAGAGGCTCCAGCTTCTGGCCCGCATCAC 949
QY 608 CGAGACCTGGAGAGGCCCAACCGCATCTCGAGAGAGCGCTGCTCCATGTTTCAACCACTGCTG 667
DB 950 CAAAAAGCT-----CGACAAAGCCCTGCGCCCAAGTTCACTCTCTGCTG 991
QY 668 CCCCCGCTGATCCGCTATCTGAGAAATCTTAAACCGGACCTGATCCCTACGTAGAGAG 727
DB 992 TCCCGCTGGCAGAAAGTACGCGAGACCTTCTAACCCGAGCTGCTGCGGCACTTCTCTC 1051
QY 728 CTGCAAGAGCCCGCAGATGATGCTGCGCGCATGCTCAAGTCTTAAGCGGAAAGAA 787
DB 1052 GTGCAAGTGGCCCATCGGATGATGAGTCCCTGCGCAAGACCTACGCGGCGCAAGAGAGCT 1111
QY 788 GGGCATCCGCGCAAGAGCATGATGATGCTGCTCATGATGCTGCTGCAACGCGCAAGAGCT 847
DB 1112 CGGCTACGAGCCCAAGCAGATCTACACCGTCTTCCATCATGCTGCTGCAACGCGCAAGAAATT 1171
QY 848 GAGGCTGACCGCACTGCTTCTGTGAGACGCGGACCCCAACCTGCGCGCAAGCTGACCA 907
DB 1172 CGAAGGCTGCGCGCGGAGAT-----GAGCGCGAGCGGCTTCCGCAATCGACGC 1222
QY 908 CGTCATCAACACCGCTGAGCTGGGCAATCTTTCAAGAGCGCGCATCAACCTGCGCGA 967
DB 1223 CACCATCAACACCGCGAGGCTGCTCATGATGAAAGGCGCGCATCGACCTGCGCGA 1282
QY 968 GCTGCGCGAGGCGAGTGGCAATCAATGAGGCGTGGGCTGCGGCGCGCGGCTGCTT 1027
DB 1283 GATGCGCAATGGCAAGGCGAGCGGCTCATGAGGAGATCAACGCGCGCGCGACATCTT 1342
QY 1028 CGGACCAACCGGCGCTGTCTATGAGGCGGCGCTGCGCAACGCGCTTATGAGCTGTTCA 1087
DB 1343 CGGCGTTTCCGCGCGCTCATGAGAGCGGCGCTGCGCTTCCCTACAGGCGCTGACCA 1402
QY 1088 CACGCGCGCTGCGCGCTGAGCGCTGAGCGAGGCGCGGATGAGCGGCTCAAGAGAGAC 1147
DB 1403 GAAGCGCGCGCGAGAGCTGAGCTTCAAGCGCGCGGCTTAAAGGCACTCAAGAGAGC 1462
QY 1148 CAACATCAACATGCTGCTGCGCGCGCGGCTCAAGTGTAGAGCTGTAAGCAACCGCGC 1207
DB 1463 CACCATCAACATG-----GCGGCAACGAGCTCAAGTGTAGCGCTGTAAGCGCGC 1513
QY 1208 CGCGCGCGCGCGAGGCGCGCGCGACCGGACCGCGCGCGCTGCGCTGAGAGCGCGG 1267
DB 1514 CAAGAACTTGGCAAGGCTTGGCAAGAGGTGAAGGCGCGGAAATCGCCTTACCACTTCA 1573
QY 1268 CGGCGCTTGCAGGAGGAGAGCGAGGCGCGGATCACACTGCGGCTGAGCGGCGCA 1327
DB 1574 CGAGTTTATGCTGCTGCGCGAGGCGGCTGCTGATGAGGCGCGGCGGCGGACCATGCTG 1633
QY 1328 CGGCGCTGCGGAGCGCAAGAGCTGATCAACCAAGATGACAGC 1369
DB 1634 CGTCTCGAATCCATGAAACGAGCAACAGAGATTCTAGCG 1675

```

RESULT 10
DVUHYDFE 1964 bp DNA linear BCT 26-APR-1993
LOCUS D.vulgaris (Fel)-hydrogenase alpha and beta subunit genes, complete cds.

ACCESSION M27212.1 GI:145098
VERSION hydrogense.
KEYWORDS D.vulgaris; oxamticus (strain Monticello) DNA; clone pJS1130.
SOURCE Desulfotribio vulgaris
ORGANISM Bacteria; Proteobacteria; delta subdivision; Desulfotribionaceae; Desulfotribio.

REFERENCE 1 (bases 1 to 1664)
AUTHORS Voordouw, G., Strang, J.D. and Wilson, F.R.
TITLE Organization of the genes encoding [Fe] hydrogenase in

```

JOURNAL Desulfotribio vulgaris subsp. oxamticus Monticello
MEDLINE J. Bacteriol. 171 (7), 3881-3889 (1989)
PUBMED 89291738
COMMENT Draft entry and printed copy of sequence for [1] kindly provided by G.Voordouw, 07-SEP-1989.
FEATURES
    source
        1..1964
            /organism="Desulfotribio vulgaris"
            /db_xref="taxon:881"
            /note="[Fe]-hydrogenase alpha subunit"
            /codon_start=1
            /transl_table=1
            /protein_id="AAA23373.1"
            /db_xref="GI:145098"
            /translation="MSRIEMEXIFYEDHAPPEKADPDKLFIQIDESCICGDCSQY
            CPTGAFDPTGDAKXIPHEBLCINCGOCLTPVPAIVESQSWTEIEKKIKADKV
            IAMPAPAVYALGDAFG.PVGVTITGKMFSALEKPELPHMSCKSP.IGMLTGLAKTYADR
            VORLTKLDPKPLPFTSCPGMHKXVESLAYPELPHMSCKSP.IGMLTGLAKTYADR
            MKYBRAKYVTVSIMPCTAKKYEGNPO.LMSDHRIDATITDRELATMIKAKIDFTK
            LPDGKDTLMSESTGATLFGVTGSMELALRYAIOAVTGKPKESMDRFGVGLGVK
            EATVAVGVADVAVVHARFRFDVCELVKAGKAPWHPERFMACPGCGCGGQPVMP
            GVLEADRRSTRMYAGLKKRLAMASASRA"
            1514..1888
            /note="[Fe]-hydrogenase beta subunit"
            /codon_start=1
            /transl_table=1
            /protein_id="AAA23374.1"
            /db_xref="GI:145100"
            /translation="MQIVNLTRRGLKAACVVTGAGALISIMTKAVNAKOLXDM
            DRINGVYDAKFPVRSQDVOYKIVADLEKPMHSKABQLLHTHVDSKAIERM
            KAQGVAPPRAKEFEGNTYPE"
BASE COUNT 409 a 638 c 633 g 284 t
ORIGIN 672 bp upstream of EcorI site.
Query Match 9.3%; Score 222; DB 1; Length 1964;
Best Local Similarity 54.4%; Pred. No. 1.6e-18;
Matches 537; Conservative 0; Mismatches 415; Indels 36; Gaps 3;

```



```

QY 912 ATGACCACCGTGGAGCTGGGCAACATCTTCAGAGAGCGCGGACATCACTGGCCGAGCTG 971
DB 936 CTGACACCGCTGAGCTGGGCTACATGATCAAGAAAGCGGATTCGACTTCGGAAACTC 995
QY 972 CCCGAGGCGCAGTGGGACATCCATGAGGCGTGGCTTCGGGCGCGCGCTGCTTCGGC 1031
DB 996 CCCGAGGCGCAGTGGGACATCCATGAGGCGTGGCTTCGGGCGCGCGCTGCTTCGGC 1055
QY 1032 ACCAGCGGCGTGTCAATGAGGAGCGCGCTGCGACCGGCTATGAGCTGTTCACGGGACG 1091
DB 1056 GTCAACCGGCGCTTCATGAGAGCGGCACTCCGCTTCGCTACGAAAGCCGTCACGGGACG 1115
QY 1092 CCGCTGCTCCGCGCTGAGCGCTGAGCGAGTGGCGCGGATGAGACGCGCATCAAGACCCAC 1151
DB 1116 AAGCCCGACAGCTGGGACCTTCAAGGCCGTGCGCGCTTGTATGAGCATCAAGAGGACACC 1175
QY 1152 ATGACCAATGTTGCCCGCGCGCGGCTCCAACTTTGAGAGCTGCTGAAGCAACCGCGCGCC 1211
DB 1176 GTCAACGCTG-----CGGTAACCGACGTCAGAGTCCCGCTGTGTCACGCGGCGCAAG 1226
QY 1212 GCGCGCGCGGAGCGCGCGCGGACGCGGACCGCGCGCTGGCTGGGAGCGCGCGCG 1271
DB 1227 CGGTTCAAGAGAGGCTCTGCGACGATGTGAAGCGGCGCAAGTCCGCTATCATTCGAA 1286
QY 1272 GCGTTCAACGAGCGGAGCGGAGCGGCGGATACACTGCGCGTGGCGCTGGCGCAAGCGG 1331
DB 1287 TACATGGCTGTCCCGCGCGCTGCTGTGGCGCGCGGTAGCGCGTCAATGCCGCGGTG 1346
QY 1332 CTGGGCAAGCGGCAAGAGCTGATGACCAATGACGAGCGCGGCGGAGCGCAAGTACGACTT 1391
DB 1347 CTGGAAGCCATGAGCGGACCCACCAACCGGCTTTACCGGCGGCTGAAGAGCGCTTGCC 1406
QY 1392 GTGAGATC 1400
DB 1407 ATGGCGAGC 1415

RESULT 12
CFU298227 3265 bp DNA linear PLN 10-JUL-2002
LOCUS Chloroella fusca hyda gene for Fe-hydrogenase, exons 1-5.
DEFINITION
ACCESSION AJ298227
VERSION AJ298227.1 GI:21732234
KEYWORDS Fe-hydrogenase; hyda gene.
SOURCE 'Chlorella' fusca.
ORGANISM 'Chlorella' fusca.
Eukaryote; Viridiplantae; Chlorophyta; Chlorophyceae;
Chlorococcales; Scenedesmeceae; Scenedesmus.

REFERENCE
1 Winkler, M., Heil, B., Heil, B. and Happe, T.
Isolation and molecular characterization of the [Fe]-hydrogenase
from the unicellular green alga Chlorella fusca
Biochim. Biophys. Acta 1576 (3), 330-334 (2002)
JOURNAL 22079022
MEDLINE 12084580
PUBMED 2 (bases 1 to 3265)
REFERENCE Happe, T.
AUTHORS Direct Submission
TITLE Submitted (13-DEC-2000) Happe T., Molekulare Biochemie, Botanisches
JOURNAL Institut der Universitaet Bonn, Karlrobert-Kreien-Str.13, Bonn,
NRM, 53115, GERMANY
FEATURES
source
1..3265
Location/Qualifiers
/organism="Chlorella' fusca"
/db_xref="taxon:3073"
67..2246
/gene="hyda"
join(67..363,523..618,866..1108,1364..1564,1773..2246)
/gene="hyda"
/EC_number="1.18.99.1"
/function="hydrogen metabolism"
/codon_start=1

```

```

/product="Fe-hydrogenase"
/protein_id="CA83290.1"
/db_xref="GI:21732235"
/translation="MCCPVVASRHARGARHVAVRAGPTSECDPEPTPOAKLPHWQOA
LDELAKPESRRLMAIQIASAVRAIAETIGLAPEDVITGOLVTLGRLMGFVYEDTL
FGADLTJIEBGETLHRLQDLHLEHPNKEEPLPFMTSCCPGVAVAKENPILPLS
SCSRQMWLGAIVKRYVAQOQVQPSDLCNVSVMCVKROGADREWFNTGAGLARD
VDHVTTAEVGIPIFERGIKINLEPESNFDNPIGEGTGALLPFTGCGMEALATVY
EVVTPKPMGRVDPEEVRGLGEGIKAEITLTKPEDDSPFKAFAGADGGITLKIAYANGL
GNAKTLISLSGKAKYDFIEVMACPGGCTGGGQPSRSTDXQILQKRDQAMYNLDERS
TIRRSHDNPIQALYDKFLGAPNSHKADLLHTHYVAGGIPEEK"
transit_peptide 67..129
/gene="hyda"
join(130..363,523..618,866..1108,1364..1564,1773..2243)
mat_peptide
/gene="hyda"
/product="Fe-hydrogenase"
/EC_number="1.18.99.1"
<67..363
/gene="hyda"
/number=1
intron 364..522
/gene="hyda"
/number=1
exon 523..618
/gene="hyda"
/number=2
intron 619..865
/gene="hyda"
/number=2
exon 866..1108
/gene="hyda"
/number=3
intron 1109..1363
/gene="hyda"
/number=3
exon 1364..1564
/gene="hyda"
/number=4
intron 1565..1772
/gene="hyda"
/number=4
exon 1773..2246
/gene="hyda"
/number=5

BASE COUNT 730 a 803 c 884 g 848 t
ORIGIN
Query Match 7.8%; Score 186; DB 8; Length 3265;
Best Local Similarity 71.1%; Pred. No. 4e-14;
Matches 246; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 1294 GGGGCGGCATCACACTGCGCGGTGCGGCAAGCGGCTGGCAACGCCAAGAGCTGA 1353
DB 1897 GGCAGGCACTACAGCTCAAGATTGACAGTGCAGTAGCGCTTGGCAATGCCAAGAGCTCA 1956
QY 1354 TCACCAAGATGCAAGCGCGGAGGCCAAGTACATTGAGAGATCATGCGTGGCCCG 1413
DB 1957 TCAAGAGCTGTGACAGAGGCGCAAGGCCAAGTATATTCATTGAGTATGCAATGCGCTG 2016
QY 1414 CCGGCTGTGTGGCGCGCGCGGCGGCGGCTGCCACCGCAAGGCCATCAGCAGAGC 1473
DB 2017 GTGGCTGTGATGGCGGAGGCGGTACAGCCCGGCACTAGTCAAGAGAGATCTTCAGAGC 2076
QY 1474 GGCAGGCGGCGGTGACCACTTGAGAGAGATTCACAGCTGCGCGCGGACGACGAGAAC 1533
DB 2077 GCCAGAGGCTATGTAACACTGAGAGCGAGTACATCCCGCGAGCCATGATATTAAC 2136
QY 1534 CGTTCATCCCGGAGCTGTACAGACGTAACCTCGAGAGCGCGCGGCGCACAGGCGGACG 1593
DB 2137 CATTCATCCAGGCGGTGTAGCAAGTCTCTAGGCGGACCAACAGCACAAGGCGACATG 2196
QY 1594 AGCTGTGACACCACTACGTGCGCGGCGGCGGTGAGAGAGAAAGA 1639

```


Db 2197 ATCTGTCGACACACTATGTGGCAGCTGGAATTCAGAGGAGAA 2242

RESULT 13
AF120457 2585 bp DNA linear BCT 25-APR-2000
LOCUS Megaspinaera elsdeni putative ABC transporter (abc) gene, partial
DEFINITION cde; and Fe-only hydrogenase (hyda) gene, complete cds.
ACCESSION AF120457
VERSION AF120457.1 GI:6650983
KEYWORDS
SOURCE Megaspinaera elsdeni.
ORGANISM Megaspinaera elsdeni.
Bacteria; Firmicutes; Clostridia; Clostridiales;
Acidimicrococcales; Megaspinaera.
REFERENCE 1 (bases 1 to 2585)
AUTHORS Acta, M. and Meyer, J.
TITLE Characterization of the gene encoding the [Fe]-hydrogenase from
Megaspinaera elsdeni
JOURNAL Biochim. Biophys. Acta 1476 (2), 368-371 (2000)
MEDLINE 20135753
PUBMED 10669801
2 (bases 1 to 2585)
AUTHORS Acta, M. and Meyer, J.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-1999) Departement de Biologie Molculaire et
Structurale, CEA-Grenoble, 17 rue des Martyrs, Grenoble 38054,
France

FEATURES
source . location/Qualifiers
1..2585
/organism="Megaspinaera elsdeni"
/strain="ATCC25940"
/db_xref="ATCC:25940"
/db_xref="taxon:907"
1..817
/gene="abc"
1..817
/gene="abc"
/codon_start=2
/transl_table=11
/product="putative ABC transporter"
/protein_id="AAF22113.1"
/db_xref="GI:6650984"
/translation="STLKVLAGOLAPTAGVTIGTGVFGEFTOOLPQDEGMEVID
VYQEHGFTVNOGRIATRLDQFLFTBEMQYTIATLGSERRRLTYLLMDPR
NYLLDDEPNTDIPRLTYLBYDTPQGVVAVSGDRVFLDRVYDKLTLEKGMWR
FYGDYSEYDEKTSYHQQEKGAAPASITKTNTIPAKTVAGLTSRSEELRITIEL
PRYEMLKGINAIAAGSNYEVSLSEQKEKVDALTEKCELEDKA"
829..869
998..2466
/gene="hyda"
988..1002
/gene="hyda"
1012..2466
/gene="hyda"
/codon_start=1
/transl_table=11
/product="Fe-only hydrogenase"
/protein_id="AAF22114.1"
/db_xref="GI:6650985"
/translation="MPEFHSREKIDIRVPIDEHNCAVQDVTKCKNCTLCRRACADT
QVLVLYSLSTGDMPICVHCGCSGACPFGAIVEVDVDAKAAKADBEKIVIPOTA
PAVRVGLGAEFGMDPGTFVGGKVAALRTLGADVDDPGADLTMEBETLHRLQ
SEELIPORTSCPMVERFAETFPYDILQHSSTSPISLIPVITYAOKNDIPK
KIVNVCVPTCAKKAIRREPELSAGGLFMBEIRDTDICTTRELAMQIDENIDFA
SLDESKFDPAGSAGGGRIFGNSGVMAAIRTAHMTGRPAPDPIPEPVNGLQ
GVKATVIFGHVFLVHAALISGLGNARAFIDLIKDAPEDVGFIEVMAFGGCGIGGG
GPKVKYQVKEARTASIKSDEETDIKASMONPEIRTIYEAFLDEBLSMAEFTL
HTYFSDKSDOLBKMNLQTNPMSPKYMPPEB"
2502..2528
stem_loop 663 a 740 c 657 g 525 c
BASE COUNT
ORIGIN

Query Match 7.5%; Score 180.8; DB 1; Length 2585;
Best Local Similarity 56.6%; Pred. No. 1.9e-13;
Matches 397; Conservative 0; Mismatches 272; Indels 33; Gaps 2;

QY 411 AAGCAGCTGTGGCTCAGAGTGGCTCCGGCCGTGGTGTGCTATGTGCGGAGACCTGGGC 470
DB 1294 AAGATGTCATCTTCAACACGGCTCCGGCTGCTCGGCTTGAAGGACCTTCGGC 1353
QY 471 CTGGCGCGGGCGGACACCCCAAGCAGCTGCGCGGACCTCCGCGCTT 530
DB 1354 ATGAGACCCCGGACCTTGTGGAAGAAAGATGTGCGAGCCCTCGTACCTGGTGG 1413
QY 531 GACGAGTGTGTGACACGCTGTTGGCCGCACTGACCATCATGAGAGGCGAG 590
DB 1414 GACTATGCTTCTGATACCGATTCGCTCCGACCTGACCATCATGAGAGGCGAGAA 1473
QY 591 CTGCTGCAACCGCTCAGCAGACCTGAGAGGCCACCGGCACTCCGAGACCGGTGCC 650
DB 1474 CTCTGCAACCGCTGCAATCGA-----AGAAATCCTTATTCG 1512
QY 651 ATGTCACAGCTGCTGCCGCTGATGCTGATGAGAAATCTTACCGGACCTG 710
DB 1513 CAGTTCACAGCTGCTGCCGCTGATGCTGATGAGAAATCTTATCCGACCTG 1572
QY 711 ATCCCTACGTGACGAGCTGCAAGAGCCCGCATGATGCTGCGCGCATGCTCAAGTCC 770
DB 1573 CTCCAGCATCTGTCTTCGACGAGAGCCCATGATGCTTCAGCCCGGTCAAGACC 1632
QY 771 TACCTAGCGGAAAGAGGATCGCCGCAAGACATGCTATGCTATCATGCCC 830
DB 1633 TATTTGCGCCAGCAAAACATGATCCCAAGAAATGTCMAAGTCTGCTTACGCCG 1692
QY 831 TGCAGCGCCAGCAGCTCGAGGCTGACC-----GGACTGCTTGTGTGAC 878
DB 1693 TGCAAGCGGAGAAAGCCGAAATCGCCGCTCTGTAAGTGAAGCTTCGGCTTTCG 1752
QY 879 GCCGACCCACCTCGCCGCTGAGTGCATGATGATGATGATGATGATGATGATGATG 938
DB 1753 GATGAACCGGAAATCGGATACCATGATGATGATGATGATGATGATGATGATGATG 1812
QY 939 TTCAAGAGCGCGGATCACTGCGGACGCTGCGGAGGAGATGGAGCATCAATG 998
DB 1813 ATCCAGGAGCAAAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1872
QY 999 GCGCGGCGCTGCGGCGCGCGGCTGCTGTTGGCAACACCGGCGGTGATGAGGCGGG 1058
DB 1873 GCGGAGCGCTCGGCGCGCGCGGTATCTTCGCACTCCGGCGGTGATGAGGCGGT 1932
QY 1059 CTGCGCAGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1100
DB 1933 ATCCGACCGGATACCATGATGATGATGATGATGATGATGATGATGATGATGATG 1974

RESULT 14
AE001705 15158 bp DNA linear BCT 02-JUN-1999
LOCUS Thermotoga maritima section 17 of 136 of the complete genome.
DEFINITION AE001705 AE000512
ACCESSION AE001705
VERSION AE001705.1 GI:4980694
KEYWORDS
SOURCE Thermotoga maritima.
ORGANISM Thermotoga maritima.
Bacteria; Thermotogae; Thermotogales; Thermotogaceae;
Thermotoga.
REFERENCE 1 (bases 1 to 15158)
AUTHORS Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L., Dodson, R.J.,
Hart, D.H., Hickey, E.K., Peterson, J.D., Nelson, W.C., Ketchum, K.A.,
McDonald, L., Utterback, T.R., Malek, J.A., Linher, K.D., Garrett, M.M.,
Stewart, A.M., Cockton, M.D., Pratt, M.S., Phillips, C.A.,
Richardson, D., Heidelberg, J., Sutton, G.G., Fleischmann, R.D.,
Eisen, J.A., Fraser, C.M. et al.
TITLE Evidence for lateral gene transfer between Archaea and bacteria
from genome sequence of Thermotoga maritima

JOURNAL Nature 399 (6734), 323-329 (1999)

MEDLINE 99287316

PUBMED 10360571

REFERENCE 2 (bases 1 to 15158)

AUTHORS Nelson, K.E., Clayton, R.A., Gill, S.R., Gwin, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Nelson, M.C., Ketchum, K.A., McDonald, L., McDaniel, L., McDaniel, T.R., Malek, J.A., Linder, K.D., Garrett, M.M., Stewart, A.M., Cotton, M.D., Pratt, M.S., Phillips, C.A., Richardson, D., Heidelberg, J., Sutton, G.G., Fleischmann, R.D., White, O., Salzberg, S.L., Smith, H.O., Venter, J.C. and Fraser, C.M.

TITLE Direct Submission

JOURNAL Submitted (01-JUN-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

FEATURES location/Qualifiers

source 1. 15158

organism="Thermotoga maritima"

gene /db_xref="taxon:2336"

CDS 139..2601

gene /gene="TM0198"

139..2601

/gene="TM0198"

/note="similar to PID:1001492 percent identity: 74.36; identified by sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="ATP-dependent Clp protease, ATPase subunit"

/protein_id="AAD35290.1"

/db_xref="GI:4980695"

/translation="MFKPSEKTAQIPFTAOEAKELGHSYVTEHLLAILKVDNSP
AVELLEWGAQSKYRSSEITISWNGKMGFVPMQTPPRARVETLEAKILGSDK
INPEHLGLDLEGGIAIHILRLKGVDAITREIIDIYSNNKLVEEEDETDY
RVEGKQEGFVNLTELAKEKLDPIGDEEIEBMQVLVARKKNVLLIGDGVKT
ALVEGAKRIVAGDVPEILKKNVIFSLDVAALVAGTKVGEFEKMKKLLQIYTKN
ILFIDEITIEIAGSAGAIIDANILKPAARBEISCIATTPDEYRRIYIDALE
RFQKIVYEPTFEETLEILGLKRYKAEHKKVITDQKALEAVYLSKRTIYHPD
KVIDIDEGAPARLKVFVLPBEKMDKLEIRSRKELAVANODEKRAQKSEEM
ELEAVRRKRYEMRRRATVAVKVDVDAEVSWSMTGVPGRKIEVEKELNLEEA
LHQRIVAODEAKAVAIRARSGLDKPPRPVGLFPTGKTELEKALAEVLF
GERALIRFDMSEYMERFSVRLGAPGVYEGEGTLEKVARFVSYLIDEXEK
APPDVFNLIQIMDGRLLTDSQREVDPRNTIIMTNSIGSSYINSKRTILGVDNN
EKEFEFKIDVLEVEKTFPPEELRIDETIIPHLKKEHIGIIDIILDRKLS
EKMKLVLTAKSEKFLVEKGFDPVYGARPLKRAIQRYVEDPLSEELRGKFNEDTIV
CRARDALRFRKGEKKEKVQ"

gene 2598..3920

/gene="TM0199"

2598..3920

/gene="TM0199"

/note="similar to SP:P24554 GB:X63155 PID:537229
PID:581233 GB:U00096 percent identity: 63.78; identified
by sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="DNA repair protein"

/protein_id="AAD35291.1"

/db_xref="GI:4980696"

/translation="MKKPVCSNGCYVSPKMGRCPCQGEYDTAEVQRRKREGESL
FLNLEAGEISLESLTSGFSLDPAFGGVPQVILSGEPGISTALQIAEFPA
SIQVFSDDLGSFGSISQVKNVMTKTIIDAKRGVPLVAGHVTKEGSIAPKVEH
MVDVAFEGEGRRTGLRLKITKNRFGPSPDEVAFELKENGFPQVENPFTGKDLV
GNVLVEFGTKFPVQIOLVSKNRTFSKRVCKGVGNRVNRLIAVISKYLLPIE
THDVYVNVVGLRTDPAADLAIALTSVSSYLEVSLHNTAAVEGIGLDRVAKYNNIN
RNLNLSKSGRIIPPIEEEGKGFVRDLKEAVSIIGGELIGTPOAD"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent identity: 65.12; identified by
sequence similarity; putative"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

gene 3883..4956

/gene="TM0200"

3883..4956

/gene="TM0200"

/note="similar to GB:D26185 SP:P37573 PID:467476
GB:AL009126 percent

```

/protein_id="AAD35296.1"
/db_xref="GI:4980701"
/translation="MEVEVILKVEGLRKDPNGVKVLENNGFSVGKGERVALLGSGSC
LKTFTLRVSGEDYOGKRVFTDKIGYVFOEPRLLPMTITENMLLIRIDRLASL
LEKVELKGFENHPMQLSGMKORVNFVAVLDPDLLDDPEPALDKTKMKDRL
LYDLNOKRFSIVFTVHNKAVFLSERIFLISGRPSKILDEVKEXAIDTDEKLF
RLEKVIERLNLRL"
/complement(9137..11479)
/gene="TM0205"
/complement(9137..11479)
/notes="similar to GP:2633959 percent identity: 63.64;
identified by sequence similarity; putative"
/codon_start=1
/translation_table=11
/product="ATP-dependent DNA helicase"
/protein_id="AAD35297.1"
/db_xref="GI:4980702"
/translation="MLCSRYFTSSLFMGEBALPTLLEFLNEVEKMLKNQVTRRIHQ
LKELEDDPLENKLDEKLAFLDYKEIPNLPFARKRYIQKLEIMIKLSMFLID
VLESGEVDLSTDIYAKGVGNPKKKLKGIELRLDLEFFPEDEPRKIFKLN
DLPGKVTTOGKIVSVETKRNKNMILTAVIDGLVHYPLKMFNDVLOTYKLOLG
KEVFTGYKSNATGYOYEIHNABEVPKSGEYRRLIPYIRLISGISOAKRIEEN
ISLCCSLKELPERILEKRLGVKDATYGMFPPTFYHLEKARERLAYEBVLQ
AFQKIRKEREKIGGIPKKEGLAEFFISLPEKLTNAQKRAHQEIRNMISEKPNR
LLOGDVSGKTVAQAQIADNENAFQIAFMPTSLIAOHYRITVESFSKNIHAL
LIGATPSEKIKSGLRNGQIDVIGTALIOEDVHFKNLGVIIDBOHRFGVORE
ALMNGKMDTLVMSATPIPRSMALAFYDGLDVTIDEMPRKEVOTMLVMDRNE
VIEFTVQEVNRGQATIVPLIESPKLVAKSVENYELSKPEPFKLMHGRIS
OEKDVMLPEAGRYDILVSTVIEVGDIVPANVAVENPERFELADHOLMGVSG
RGQEAIVCFPLVGDVGEAMERLRFELNTDGLKLAEDYLKTRGPGEFVGVOHGLSG
FKVADYRLDKLEMARDEVOIEVIEGLPEIKLEIVG"
/complement(11527..12042)
/gene="TM0206"
/complement(11527..12042)
/gene="TM0206"

Query Match      6.9%; Score 166.6; DB 1; Length 15158;
Best Local Similarity 54.1%; Pred. No. 7.2e-12;
Matches 402; Conservative 0; Mismatches 314; Indels 27; Gaps 2;

```

```

QY 891 CTCGCCAGCTGGACCAAGCTATCAACACCGTGAAGCTGGGCAACATCTTCAAGAGCCG 950
DB 6183 -----GTTGATGTTGTTCTTCAACCAAGAGGAATTGGGAAGCTATAAGATGAAG 6234
QY 951 GGATCAACCTGGCCGAGCTGCCAGAGCCAGTGGGCAATTCATAGGCGCTGGGCTGG 1010
DB 6235 AAGATACCGTTTGGCAACCTTTCAGAGGAAGATACGACGCCCTTGAATCTTTAGG 6294
QY 1011 GGCGCGGCGCTGCTGTTGGGACACACCGCGGCTGCATGAGAGCGCGCTGGCAGCGCC 1070
DB 6295 GGAGCGCGCGCACTCTTGGCCTTACAGCGCGTGTGATGAGAGCTGCTTGAACACTGGC 6354
QY 1071 TATGACCTGTTCACGGGCAAGCGCGCTGCCGCTAGCCTGAGAGGAGTGGCGGCATG 1130
DB 6355 TACGAACCTCAAAACGGGAAGGCACTACCGAAGATCGTTTGAAGAGGTGAGAGGACTC 6414
QY 1131 GACGCGATTAAGAGACCAACT 1153
DB 6415 AAGAGTGTCAAGGAAGCCGAAT 6437

RESULT 15
DHYDC
LOCUS D.vulgaris hydC gene for hyd gamma periplasmic hydrogenase.
DEFINITION X57838.1 GI:40827
ACCESSION X57838
VERSION 1
KEYWORDS hydC gene; periplasmic Fe-hydrogenase.
SOURCE Desulfovibrio vulgaris
ORGANISM Bacteria; Proteobacteria; delta subdivision; Desulfovibrionaceae;
Desulfovibrio.
REFERENCE 1 (bases 1 to 1940)
AUTHORS Stockermans J.P.W.G.
TITLE Direct Subdivision
JOURNAL Submitted (15-JUN-1991) J.P.W.G. Stockermans, Agricultural Univ of
Wageningen, Dept of Biochemistry, Dreijenlaan 3, 6703 HA
Wageningen, The Netherlands
REFERENCE 2 (bases 1 to 1940)
AUTHORS Stockermans J., van Dongen W., Kaan A., Van Den Berg W., and
Veeger C.
TITLE Hyd Gamma, a gene from Desulfovibrio vulgaris (Hildenborough)
JOURNAL encodes a polypeptide homologous to the periplasmic hydrogenase
COMMENT FEMS Microbiol. Lett. 58, 217-222 (1989)
FEATURES
source Location/Qualifiers
1..1940
/organism="Desulfovibrio vulgaris"
/strain="Hildenborough"
/db_xref="taxon:881"
37..41
46..1918
/gene="hydC"
46..1918
/gene="hydC"
46..1918
/gene="hydC"
prim_transcript <46..1918
/gene="hydC"
46..1866
/gene="hydC"
/codon_start=1
/translation_table=11
/product="Hyd gamma"
/protein_id="CAA40970.1"
/db_xref="GI:40828"
/db_xref="SPTREMBL:Q46606"
/translation="NNAFLNGEVRCEPGRITLLEARENGHFIPTLCELADIGHAGT
CRVCLVEIRDEKAGPOIVTSCITPVEEGRIETRTPEVRMQRLQVELLLADHDHDC
AACARHGDCELDVAFQVGLSTGRHHPDYSARRDVSFVVRMGKICILRCVA
VCRNQGVDALVVTNGITETGLAHNSQSASDCVCGCGCTGVGALARGDVER
VDVYXDEPRTVTFQPARAVRGLEBERGLPRGSSVEGYPFALRLKADVDLITRA
ADVLMEBGTIELQRLRGAKPLFTSCCPGVNFAEKHLPILPHVSTTRSPQOGLG
ALAKYLLARTNNAVERMRVVSIMPTAKEEAARPEFRDQVADVAVLTTRFARL
LRREGIDLAGLEPSPCDPLMGARATGAAVIFGTGTVMEALRTVHYVHNGELAPVE

```

LHAIKGVNVEAVVPILEGNSVKVAVVHGLKARQVNEAVLAGKADHVEVMACP
GGCNDGGQPSKRAVNPNAQARRAALFSLDAENALQSHNPLIGVYESFLGEPCS
NLSHRLHTRYGDRKSEVAAYTMRDIWHMTLIGRIVRGDSB

terminator

1892..1918

/gene="hydc"

BASE COUNT 343 a 597 c 641 g 359 t

ORIGIN

Query Match 6.7%; Score 160; DB 1; Length 1940;

Best Local Similarity 52.3%; Pred. No. 7,3e-11;

Matches 451; Conservative 0; Mismatches 375; Indels 36; Gaps 3;

```
QY 374 GCGGAGGTTGCCAAGCCCAAGAGACACCCAGCGCAAGACGCTTGGCGTGCAGGTGGC 433
DB 690 CGTGAGAGGTGTCAATCGACTATCTCTAGACCCCGAAATCGTCAACCGTGTCCAGTTGCG 749
QY 434 TCCGCGCGCTTGTCTGCTATTGGCGAGACCTGGGCGCTGGCGCGCGGCGCCACCAACC 493
DB 750 CCGGCGGCTGGCGGCTGGCGCTGGCGAGAGTTCGGGCTGCTCCGGTTCAAGCGTGA 809
QY 494 CAAGCAGTGGCCGAGGGCTCCGCGCGCTTGGCTTGAAGAGTGTGACAGCTGTT 553
DB 810 AGGCGAGGTGCCACGGGCTTGGCGCTTGGCGAGACGTGTACTCGATACCAACTT 869
QY 554 TGGGCGGACCTGACCATCATGAGAGAGGAGGAGAGTGTGACCGGCTCACCGAGCA 613
DB 870 CGCAGCGACCTCGTCAATCATGAGAGAGGAGCAGCACTCTGCAACGCTCT----- 920
QY 614 CTTGAGAGGCCACCGGACTCCGAGAGCGCTGCCATGTTCAACGAGTGTGCGCCGG 673
DB 921 -----TGGGGCGGGGGCGAAGCTGCCGCTCTTCACTCTCTGCGCCGG 965
QY 674 CTGATGCTATGCTGAGAAATCTTAACCGGACCTGATCCCTACGTGACAGCTGCA 733
DB 966 CTGGGTAAATTTGCCGAGAGCACTCCCGACATCTGCCGATGTCTGACCAACAG 1025
QY 734 GAGCCCCAGATGATGCTGGGCGCATGTCAGTCTTACCTAGCGGAAAGAGGCAAT 793
DB 1026 CTGCGCTCAGCAGTGTCTTGGCGCATGGCCAGACCTTATCTTGGCGGACCAATGAACT 1085
QY 794 CGCGCCAAAGACATGTCATGTCATATGCTTGCATGCTTGCACGCGCAAGCAGTGGAGGC 853
DB 1086 CGCACCGGAGAGATGCGCTGATGATGATGCTTGCACGCGCAAGAGAGAGGC 1145
QY 854 TGACCGGACGTGTTCTGTGTGAGAGCGGACCCACCTTGGCGCAGTGTGACCACTCAT 913
DB 1146 CGCAGCGCCGAATTCAGGCGCGACGCTG-----TCGGGATGTGACCGCAGTGTCT 1196
QY 914 CACCACCGTGGAGCTGGGCAACATCTTCAAGAGCGCGGACATCACTGGCCGAGCT--- 970
DB 1197 CACCAAGCGTGAATGCTGCGCTTCTTCCGCGGTGAGGGCATAGACTTCGCGGACTCGA 1256
QY 971 GCGCGAGGCGGATGAGCAATCAATGGGCTGGGCTGGGCGCGCGGCTGTGTTGG 1030
DB 1257 ACCCTGCGCTGCGACACCCCTGATGGGCGGCGCAACCGGGCGGCTGTCACTTCGG 1316
QY 1031 CACCACCGGCGTGTCTATGAGGCGGCGCTGGCAGCGCTTATGAGCTGTTCAGGGCAC 1090
DB 1317 TAGCAGAGGCGGGGTAAAGAGCGGACATGCTTACGATGTGTGAAAGCGCAA 1376
QY 1091 GCGGCTGCGCGCTTGAAGCTTGAAGAGTGCAGGATGACGCGATCAAGAGAGCAA 1150
DB 1377 GGAAGTGCAGCCCAAGAACTGATGCCCTGGCGGATAGAGAACTGTGAGGCTGT 1436
QY 1151 CATCACCATGATGCTGCGCGCGGCTTCAAGTTGAGAGAGTGTGAGAGCAGCGCGCC 1210
DB 1437 GGTCCGCTTGTGAGAGGTAAAGGTTCCCTGGAAGGTGCGCGTGTGATGGGCTCAAGGC 1496
QY 1211 CGCGCGCGCGAGCGCGCGCG 1232
DB 1497 TGCCCGGAGATGTGAGAGCG 1518
```

Search completed: June 8, 2003, 11:53:21
Job time : 6045 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 09:13:01 ; Search time 526 Seconds
(without alignments)
10270.996 Million cell updates/sec

Title: AJ012098
Perfect score: 2399
Sequence: 1 ATCTTACATGAACACACAAA.....TGTAAGCAGCAGCTTCTGAG 2399

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N Geneseq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	3.9	1737	20	AAZ25199
2	83.2	3.5	114955	20	AA53491
3	75	3.1	109519	22	AA508693
4	72.6	3.0	114955	20	AA53491
5	71.4	3.0	1552	22	AAH13714
6	71.4	3.0	1577	22	AAH27127
7	70.2	2.9	65140	22	AAH17184
8	70.2	2.9	125401	22	AAH17186
9	69.8	2.9	1836	21	AAH70438

10	69.8	2.9	1944	21	AAA70434
11	69.8	2.9	1944	21	AAA70435
12	69.8	2.9	1944	21	AAA70437
13	69.8	2.9	2025	21	AAA70433
14	69.8	2.9	2025	21	AAA70436
15	69.8	2.9	2310	21	AAA51944
16	69.8	2.9	2316	21	AAA51945
17	69.8	2.9	2316	21	AAA51949
18	69.8	2.9	2322	21	AAA51946
19	69.8	2.9	2328	21	AAA51947
20	69.8	2.9	2334	21	AAA51948
21	69.8	2.9	2358	21	AAA70447
22	69.8	2.9	2358	21	AAA51966
23	69.8	2.9	2466	21	AAA70446
24	69.8	2.9	2547	21	AAA70445
25	69.8	2.9	4608	21	AAA70473
26	69.8	2.9	4689	21	AAA70471
27	69.8	2.9	27541	22	AAH17185
28	68.4	2.9	3009	24	ABL40024
29	68.2	2.8	53789	19	AAV21187
30	67.2	2.8	1944	21	AAA51613
31	67.2	2.8	2101	22	AAH24245
32	67.2	2.8	2113	22	AAH24245
33	67.2	2.8	2466	21	AAA51614
34	67.2	2.8	2547	21	AAA51615
35	67.2	2.8	2550	24	ABL39955
36	65.4	2.7	44377	18	AAH78508
37	65.4	2.7	44377	18	AAH80414
38	64.6	2.7	50937	21	AAH09469
39	64.6	2.7	3465	12	AAQ14478
40	63.8	2.7	3993	12	AAQ15153
41	63.6	2.7	2322	21	AAA51959
42	63.6	2.7	2341	21	AAA51951
43	63.6	2.7	3015	24	ABL39983
44	63.4	2.6	1182	22	AAH14507
45	63.4	2.6	1248	19	AAV11297

ALIGNMENTS

RESULT 1
AAZ25199 standard; DNA; 1737 BP.
XX AAZ25199;
AC 13-DEC-1999 (first entry)
XX 13-DEC-1999 (first entry)
DE Clostridium pasteurianum hydrogenase #2.
XX Clostridium; hydrogenase; blue-green alga; hydrogen; microbe;
KW Shine-Dalgarno; ss.
XX
OS Clostridium pasteurianum.
XX
PN JP11253166-A.
XX
PD 21-SEP-1999.
XX
PF 13-MAR-1998; 98JP-0062601.
XX
PR 13-MAR-1998; 98JP-0062601.
XX
(AGEN) AGENCY OF IND SCI & TECHNOLOGY.
PA (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.
XX WPI; 1999-583699/50.
XX
PT Preparation of heterologous hydrogenase in blue-green algae - useful
for production of heterologous and is environmentally friendly
XX
XX Example 4; Fig 10; 12pp; Japanese.

XX A method has been developed for the preparation of hydrogenase in which
CC a blue-green alga introduced by a recombinant vector containing a
CC heterologous hydrogenase gene is cultured in a medium to form and
CC accumulate hydrogenase in the microbe body. Also claimed are: (1) a
CC blue-green alga introduced by a recombinant vector containing a gene
CC having a Shine-Dalgarno sequence consisting of AAGGAA upstream of the
CC hydrogenase structural gene derived from a Clostridium genus microbe;
CC (2) a method for the production of hydrogen by culturing the above
CC blue-green alga under light irradiation; (3) a method for the production
CC of hydrogen by culturing the above blue-green alga anaerobically and
CC then under light irradiation; (4) a method for the production of
CC hydrogen by reacting the above blue-green alga with an electron
CC receptor; and (5) a method for the production of hydrogen by culturing
CC the above blue-green alga anaerobically and then reacting it with an
CC electron receptor. The method is mild to environment. The present
CC sequence represents a Clostridium pasteurianum hydrogenase nucleotide
CC sequence from an example from the present invention.

XX Sequence 1737 BP; 565 A; 363 C; 373 G; 436 T; 0 other;

Query Match 3.9%; Score 93; DB 20; Length 1737;

Best Local Similarity 48.1%; Pred. No. 1.1e-07;

Matches 359; Conservative 0; Mismatches 355; Indels 33; Gaps 2;

QY 398 CGACCCGACCGGACGACGCTGCGTGCAGTGGCTCCGCGCTTCGCTGCTATTC 457
DB 660 CAATGCCCGCGAAGAAACATGTGATCGCTATGCTCCGACCTCCGCGCTGACATCGG 719
QY 458 CGAGACCCCTGGCG 517
DB 720 TGAACCTTTTAAATATGAGATTTGCGCTTCACGTACAGGAAATTTATAGGCTTCGCG 779
QY 518 CCGCCTCGGCTTTCAGAGGTGTTTGCACGCTGTTTGGCGCGCGCGCGCGCGCGCGCG 577
DB 780 CCAGCTTGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 839
QY 578 GAGAGGACGAGAGCTGTGACACCGCTCAGACCGACCTTGAGAGCGCGCGCGCGCGCG 637
DB 840 AGAGGCTACAGAACTGGTTCAACGCAAT-----GGAGAAATTA 875
QY 638 CGAGCGCGCTCCGATGTTACAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 697
DB 876 TGAACCTTTCCCATGTTTAAAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 935
QY 698 TTACCCGACCTGATCCCTACGTGAGCAGCTCAAGAGCGCGCGCGCGCGCGCGCGCG 757
DB 936 TTATCCGAACTCCTAAATATCTTTTCATCAGCTAATATCACTCAACAATTTTGGTAC 995
QY 758 CATGTCAGTCTCTTACCTTACGCGGAAAGAGGCGATGCGCGCAAGACATGCTATGCT 817
DB 996 CGCTAGTAAACCTTATTTATCTTATCTGACATCAGCGGCTTTGACCGCGCGCGCGCG 1055
QY 818 GTCCATCATGCTCTGACGCGGACAGTGCAGAGCTGACCGCGAGCTGTTCTGTGTGA 877
DB 1056 GAAGTATATGCTCTGATCTCAAAAAATTTGAAGCAGATCGCGCGCAATGAAAAAGA 1115
QY 878 CGCGGACCCGACCTGCG 937
DB 1116 CG-----GCTACGCGATATCGATGCTGTTATCACCACCGAGAACTGCGCAAAAT 1166
QY 938 CTTCAAGAGCGCGGCGATCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 997
DB 1167 GATTAAGATGCTAAATATCCCTTTTGTAACTTGAAGATGGAAGCAACCTCTAT 1226
QY 998 GGGCGTGGGCTCG 1057
DB 1227 GGAAGAAATACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1286
QY 1058 GCTGCGACGCGCTATGAGCTGTTACGCGGACGCGCGCGCGCGCGCGCGCGCGCG 1117
DB 1287 TCTGCGAGTGAAGAAAGACTTTGCTGAAACGCTGAAGATTCATATATTAAGA 1346

QY 1118 GGTGCGCGCATGACGCGCATCAAGCA 1144

DB 1347 AGTTCCGCGACGATGATGATCAAGA 1373

RESULT 2

AA53491/C

ID AAX53491 standard; DNA; 114955 BP.

AC AAX53491;

DT 05-JUL-1999 (first entry)

DE Human adenosine A1 receptor antisense oligonucleotide fragment.

XX Antisense oligonucleotide; multiple target; antisense treatment;

KM impaired respiration; inflammation; lung disease;

KM pulmonary vasoconstriction; inflammation; allergic rhinitis;

KM acute asthma; allergy; asthma; impeded respiration;

KM respiratory distress syndrome; pain; cystic fibrosis;

KM pulmonary hypertension; pulmonary vasoconstriction; emphysema;

KM chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;

KM colon cancer; breast cancer; lung cancer; pancreatic cancer;

KM hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;

KM prostate cancer; 88.

XX Synthetic.

OS MO9913886-A1.

PN 25-MAR-1999.

PD 17-SEP-1998; 98WO-US19419.

PF 09-JUN-1998; 98US-0093972.

PR 17-SEP-1997; 97US-0059160.

XX (UYEC-) UNIV EAST CAROLINA.

PA NYce JW;

PI WPI; 1999-229400/19.

DR New antisense oligonucleotides used in treatment of, e.g. pulmonary

XX vasoconstriction

PT Disclosure; Page 37; 120pp; English.

PS The specification describes antisense oligonucleotides (AAX52869-X55271)

XX directed against at least 2 mRNAs selected from target genes, coding and

CC non-coding regions of RNAs corresponding to target genes, gene

CC initiation codons, genomic flanking regions, intron-exon borders, the

CC 5'-end, the 3'-end and the junction-section between coding and non-coding

CC regions and all segments of RNAs encoding proteins associated with one

CC or more diseases, conditions or mixtures. The antisense oligonucleotides

CC may be derived from sequences AAX5272-74. These multiple target

CC oligonucleotides (specifically AAX5180-211) can be used for the

CC antisense treatment of diseases and conditions. Typical diseases and

CC conditions are those associated with impaired respiration and

CC inflammation, including lung diseases, pulmonary vasoconstriction,

CC inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded

CC respiration, respiratory distress syndrome, pain, cystic fibrosis,

CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic

CC obstructive pulmonary disease (COPD), and cancers such as leukemias,

CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,

CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,

CC hepatic metastases, as well as all types of cancers which may metastasize

CC or have metastasized to the lungs, including breast and prostate cancer.

XX

SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match 3.5%; Score 83.2; DB 20; Length 114955;

Best Local Similarity 31.5%; Pred. No. 7.6e-06;

Qy	1267	GC	CG	GG	CG	CTT	CA	CC	AG	CG	AG	GA	CG	GG	CG	GC	AT	CA	CT	CG	CG	GT	GC	CT	GG	AC	CA	1326																									
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:																									
Db	103971	CCG	CA	CA	CG	AG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	103912																									
Qy	1327	AC	GG	CG	CT	GG	CA	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1386																										
Db	103911	CC	AA	GA	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	103852																										
Qy	1387	ACT	TG	TG	GA	GA	TG	AT	CT	AT	CT	AT	CT	AT	CT	AT	CT	AT	CT	AT	CT	AT	CT	AT	CT	AT	1444																										
Db	103851	G	CG	CA	GA	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	103792																										
Qy	1445	CT	CC	AC	CG	CA	CA	AG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1504																										
Db	103791	AC	GC	CG	CG	CA	VG	AC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	103732																										
Qy	1505	GT	CC	AC	CG	CT	GC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1564																										
Db	103731	CC	CG	GC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	103672																										
Qy	1565	CG	AG	AG	CG	CG	CT	GG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1620																										
Db	103671	CC	GC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	103616																										
RESULT 3																																																					
AA08693																																																					
ID	AA08693	strand; DNA; 109519 BP.																																																			
XX	AA08693;																																																				
XX	26-SEP-2001	(first entry)																																																			
DE	Micromonospora DNA encoding biosynthetic enzymes for Evernimycin.																																																				
XX	Evernimycin; antibiotic; bottle-neck gene; orthomycin;																																																				
KM	fermentation; ds.																																																				

```
FT CDS 8342..9364
FT /tag= m
FT /product= "EvdH"
FT RBS 8333..8336
FT /tag= n
FT CDS complement (9463..10224)
FT /tag= o
FT /product= "EvdI"
FT RBS complement (10232..10235)
FT /tag= p
FT CDS 10424..11176
FT /tag= q
FT /product= "EvdJ"
FT 12027..12455
FT /tag= r
FT /product= "EvdK"
FT /partial
FT /note= "No start codon"
FT CDS complement (12108..13022)
FT /tag= b
FT /product= "EvdL"
FT RBS complement (13027..13030)
FT /tag= c
FT CDS complement (14410..15363)
FT /tag= u
FT /product= "EvdA"
FT RBS complement (15369..15373)
FT /tag= v
FT CDS complement (15380..16414)
FT /tag= w
FT /product= "EvdB"
FT complement 16419..17873
FT /tag= x
FT /product= "EvdC"
FT CDS complement (17870..18934)
FT /tag= y
FT /product= "EvdD"
FT 18974..20906
FT /tag= z
FT /product= "EvdE"
FT CDS 21064..22542
FT /tag= aa
FT /product= "EvdF"
FT RBS 21056..22542
FT /tag= ab
FT 22748..24172
FT /tag= ac
FT /product= "EvdG"
FT RBS 22736..22740
FT /tag= ad
FT CDS complement (24177..25223)
FT /tag= ae
FT /product= "EvdH"
FT RBS complement (25230..25233)
FT /tag= af
FT 25550..26626
FT /tag= ag
FT /product= "EvdI"
FT CDS 26685..30479
FT /tag= ah
FT /product= "EvdJ"
FT 26672..26676
FT /tag= ai
FT CDS complement (30557..31876)
FT /tag= aj
FT /product= "EvdK"
FT RBS complement (31885..31888)
FT /tag= ak
FT CDS complement (31941..32882)
FT /tag= al
FT /product= "EvdL"
FT complement (33167..34405)
FT /tag= am

FT RBS /product= "EvdM"
FT /tag= an
FT CDS complement (34414..34418)
FT /tag= ao
FT /product= "EvdN"
FT RBS complement (35219..35221)
FT /tag= ap
FT CDS complement (35294..36238)
FT /tag= aq
FT /product= "EvdO"
FT CDS complement (36235..36963)
FT /tag= ar
FT /product= "EvdP"
FT CDS complement (36998..38026)
FT /tag= as
FT /product= "EvdQ"
FT CDS complement (38072..38566)
FT /tag= at
FT /product= "EvdR"
FT CDS complement (38892..40163)
FT /tag= au
FT /product= "EvdS"
FT CDS complement (40216..40890)
FT /tag= av
FT /product= "EvdT"
FT RBS complement (40899..40902)
FT /tag= aw
FT CDS complement (40887..41576)
FT /tag= ax
FT /product= "EvdU"
FT CDS complement (41679..42707)
FT /tag= ay
FT /product= "EvdV"
FT RBS complement (42714..42717)
FT /tag= az
FT CDS complement (42810..43799)
FT /tag= ba
FT /product= "EvdW"
FT RBS complement (43807..43811)
FT /tag= bb
FT CDS complement (43799..44866)
FT /tag= bc
FT /product= "EvdX"
FT CDS complement (45014..45760)
FT /tag= bd
FT /product= "EvdY"
FT RBS complement (45767..45770)
FT /tag= be
FT CDS complement (45962..46714)
FT /tag= bf
FT /product= "EvdZ"
FT RBS complement (45952..45956)
FT /tag= bg
FT CDS complement (47156..49234)
FT /tag= bh
FT /product= "EvdA"
FT CDS 51627..52715
FT /tag= bi
FT /product= "EvdB"
FT 51629..51622
FT /tag= bj
FT CDS 52889..53557
FT /tag= bk
FT /product= "EvdC"
FT 53554..54207
FT /tag= bl
FT /product= "EvdA"
FT CDS complement (54362..55117)
FT /tag= bm
FT /product= "EvdB"
FT RBS complement (55125..55128)
FT /tag= bn
```

FT CDS complement (55135..56094)
 FT /*tag= do
 FT /product= "EybC"
 FT RBS complement (56100..56103)
 FT /*tag= bp
 FT CDS complement (56184..56813)
 FT /*tag= bq
 FT /product= "EybC2"
 FT CDS 56961..58709

Query Match 3.1%; Score 75; DB 22; Length 109519;
 Best Local Similarity 45.2%; Pred. No. 0.00022;
 Matches 463; Conservative 0; Mismatches 535; Indels 27; Gaps 4;

QY 320 CGCGCTGCGCGACCCGCTGCGAGCGCCTTTGAGCATGTCCAGAGCGCTCCGCGA 379
 DB 78769 CGCGCGCTGCTATCAAGAGCCATGGCGGACACACACTGACTCTGACACACAGCGCGA 78828
 QY 380 GCTTGCCAAAGCCCAAGAGACCCCAAGCGCAAGCATCTGCTGAGGTGCTCCGCG 439
 DB 78829 TCTGAGACGACTTCCGACCGGCGCGGCTGCTGACGACGACCCGACCGGCGAC 78888
 QY 440 CGTCTGTGCTATTGCGAGACCTGAGCTGCGGCGCGGCGCGGCGCGGCGCGGCGG 499
 DB 78889 CCGCGACGAGCTGAGAGCGCTGCTGCAAGGCTGCGCGCGGCGCGGCGCGGCGGCGG 78948
 QY 500 GCTGCGCGAGGCGCTCCGCGCTGCTGCAAGAGTGTGTCACGCTGTTGACGCGCTTTGGCGG 559
 DB 78949 CCGGTTCCGCGGCGCACTGACCTTCCGACCGCGGCGCTGCGGCGCGGCGCGGCGCGG 79008
 QY 560 CGACCTGACCATGAGAGGAGGCGAGCGAGCTGCTGACCGCTCAGCGAGCACTGGA 619
 DB 79009 CCGCAACGGGATGACCTCGCGCGGTCACCAAGCGCGCGGCGGCGGCTGCTGCGCTGCT 79068
 QY 620 GCGCCACCGGCACTCCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 679
 DB 79069 CCGCGCGCGAGAGCGCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 79128
 QY 680 CGCTATGCTGGAAGAAATCTTACCGGAGCTGATCCCTAGTGAAGAGCTGCAAGAGCGG 739
 DB 79129 GAGGTT-----CGCCGAGCGCACCGCCAGTGGCCACCGGCGGCGGCGGCGGCGG 79173
 QY 740 CCAGATGATGCTGCGCGGCGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799
 DB 79174 CCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 79233
 QY 800 AAAGACATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 859
 DB 79234 GCTCGACGCGCGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 79293
 QY 860 CGACTGCT 919
 DB 79294 CTTCAGAGTTTACTCGGCGCGGAGCTCGGCG--GGGAGCTGCGCGCGGCGGCGGCGGAGAT 79350
 QY 920 CGTGAGCTGCGCAACTCTTCAAGAGCGCGGCGCATCACTGCGCGAGCTGCGCGGAGG 979
 DB 79351 CGTGGCGCGCGCGAGACCGGACATCAAGGCGCGCATCGGCGGCTGCGCGCGGCGGCGG 79410
 QY 980 CGAGTGGAGCAATC-----AATGGCGCTGCGGCTGCGGCGCGGCGGCTGCTGCTGCGGCG 1033
 DB 79411 CGTACCGCTGCGCGCGCGCGCGGCGGAGTCTGCTGCGGAGCACTGCTGCTGCTGCTGCTGCTG 79470
 QY 1034 CACCGCGGCT 1093
 DB 79471 CCGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 79530
 QY 1094 GCTGCGCGCGCTGAGCTGAGCGAGGCTGCGGCGCATGAGCGCATCAAGAGCAACAT 1153
 DB 79531 GCTGACGCGGCTGCGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 79590
 QY 1154 CAC---CATGCTGCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1210
 DB 79591 CCGCGGCT 79650

QY 1211 CGCGCGCGCGGAGCGCGCGCGGACGCGACCCCGGCGCGCTGCGGAGCGGCGCGG 1270
 DB 79651 CAACCGGAGGAGCGCGGCGCGGCTGAGCTCTGCTGCTGCTGCGGAGCGACCGGCGG 79710
 QY 1271 GCGCTTACCAAGGAGAGCGGAGGCGCGCATCACTGCGGCTGCGGCGGCGGCGG 1330
 DB 79711 GGAACCTGGGATGCGCAAGACCGCGGAGCCGCTGCGGCTGCGGCGGCGGCGG 79770
 QY 1331 GCTGG 1335
 DB 79771 CCGGG 79775

RESULT 4
 AAX53491
 ID AAX53491 standard; DNM; 114955 BP.
 XX AAX53491;
 AC
 XX
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE Human adenovine A1 receptor antisense oligonucleotide fragment.
 XX
 XX Antisense oligonucleotide; multiple target; antisense treatment;
 KM impaired respiration; inflammation; lung disease;
 KM pulmonary vasoconstriction; inflammation; allergic rhinitis;
 KM acute asthma; allergy; asthma; impeded respiration;
 KM respiratory distress syndrome; pain; cystic fibrosis;
 KM pulmonary hypertension; pulmonary vasoconstriction; emphysema;
 KM chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
 KM colon cancer; breast cancer; lung cancer; pancreatic cancer;
 KM hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
 KM prostate cancer; ss.
 XX
 XX Synthetic.
 OS
 XX
 PN W09J13886-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 17-SEP-1998; 98WO-US19419.
 XX
 PR 09-JUN-1998; 98US-0093972.
 PR 17-SEP-1997; 97US-0059160.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 XX
 PI Nyce JW;
 XX
 DR WPI; 1999-229400/19.
 XX
 PT New antisense oligonucleotides used in treatment of, e.g. pulmonary
 PT vasoconstriction
 XX
 PS Disclosure; Page 37; 120pp; English.
 XX
 CC The specification describes antisense oligonucleotides (AAX52869-X55271)
 CC directed against at least 2 mRNAs selected from target genes, coding and
 CC non-coding regions of RNAs corresponding to target genes, gene
 CC initiation codons, genomic flanking regions, intron-exon borders, the
 CC 5'-end, the 3'-end and the juxta-section between coding and non-coding
 CC regions and all segments of RNAs encoding proteins associated with one
 CC or more diseases, conditions or mixtures. The antisense oligonucleotides
 CC may be derived from sequences AAX5272-74. These multiple target
 CC oligonucleotides (specifically AAX5180-271) can be used for the
 CC antisense treatment of diseases and conditions. Typical diseases and
 CC conditions are those associated with impaired respiration and
 CC inflammation, including lung diseases, pulmonary vasoconstriction,
 CC inflammation, allergic rhinitis, acute asthma, allergy, asthma, impeded
 CC respiration, respiratory distress syndrome, pain, cystic fibrosis,
 CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic
 CC obstructive pulmonary disease (COPD), and cancers such as leukemias,

CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer, pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as well as all types of cancers which may metastasize or have metastasized to the lungs, including breast and prostate cancer.

Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match	3.0%;	Score 72.6;	DB 20;	Length 114955;
-------------	-------	-------------	--------	----------------

Best Local Similarity 32.5%; Pred. No. 0.00058;

Matches 411; Conservative 98; Mismatches 742; Indels 14; Gaps 4;

OY		201	CGCGCAGCTCTTGACGAGCGCGGCAGGTGCCCCCGCGCTCCGTCGCAACGACACC	260
Db		104391	CCBGGCGCGCGCGCGGCCGGSNNNDNCCGBGCGCBGGGCGCGCGCGCCSNNND	104450
OY		261	GNGCGTGTAGCCCTTGACAACATTAGAGCGCCCGCACGCGCCCTAGGCAACGCTTGTC	320
Db		104451	NNCCCGBGCCBGGGCGCGCCCGCCCGCSNNNDNCCGBGGCCBGGGCGCGCCCGGC	104510
OY		321	GGCGTTGCCGACCCGCTGCGAGGCGCTTTGAATCATGTCACAGGCGCTCGCGAG	380
Db		104511	SNNNDNCCGCBGGGCGCGCGCGCSNNNDNCCGBGGCCBGGGCGCGCGCC	104570
OY		381	CTTGCCAAAGCCCAAAGACACCCACGCCAAGCAACTGTGCTGAGGTGCTCCGGCC	440
Db		104571	GSNNNDNCCGBGGGCGCGCGCGCSNNNDNCCGBGGCCBGGGCGCGCGCSN	104630
OY		441	GTTTCG---TGTCGCTATTGCGAGACCTCTGGGCTGGCGCGCGGCCAACCCCAC	496
Db		104631	NNDNCCGCBGGGCGCGCGCGCSNNNDNCCGBGGCCBGGGCGCGCCSNNNDNC	104690
OY		497	GCAGCTGGCCGAGGCGCTTCGCGCGCTCGCTTTGACGAGGTGTTGAACGCTGTTGG	556
Db		104691	CBGGCCBBGGGCGCGCSNNNDNCCGCBGGGCGCGSGSNNNDNCCGCBGGCCBGG	104750
OY		557	CGCC---GACCTGACCATATGAGAGGAGGACGAGCTGCGACCGCTCACGAGCA	613
Db		104751	CGCSNNNDNCCGCBGGCCBGGCGGSNNNDNCCGBGGCCBGGGCSNNNDNCCBGG	104810
OY		614	CCTGAGAGGCCCAACCGCACTCCGACGAGCGCGCTGCCATGTTACACGCTGCGCCCG	673
Db		104811	CCBGGGSNNNDNCCGCBGGCCBGGGSNNNDNCCGBGGCCBBSNNNDNCCGCBGGCCB	104870
OY		674	CTGGAATCGCTTAGCTGGAATAATCTTAACCGGACCTGATTCCTTACGTGACGCTGCAA	733
Db		104871	SNNNDNCCGCBGGCCSNNNDNCCGBGGCCBGGGCGCGCCCGCGSCGSCGCSNNNDN	104930
OY		734	GAGCCCCAGATGATCTGCGCGGCATGTGCAAGTCTCTACCTGAGAAAGAGGGCAT	793
Db		104931	NCCBGGCCBGGGCGCGCCCGCGCGCGCSNNNDNCCBGGCCBGGGCGCGCCCGCG	104990
OY		794	CGCGCCAAAGACATGTCATGTGTCATCATGCTCGACGCGCAAGCATCGGAGCG	853
Db		104991	CGGCGCGSNNNDNBGGCCBGGGCGCGCGCGCGCGCGGCG--CGSNNNDNCGCBGGCG	105049
OY		854	TGACCGCGACTGTTTGTGTGGAACCCGACCCACCCCTGCGCGCACTGACACACTCAT	913
Db		105050	CGCCGCGCGCGCGCGCSNNNDNCCBGGGCGCGCGCGCGCGCGCGCSNNNDNCB	105109
OY		914	CACACACTGAGACTGGGCAACATCTTCAAGAGAGCGCGCATCAACTGCGAGCTGCG	973
Db		105110	GGGCGCGCGCGCGCGCGCSNNNDNCCBGGGCGCGCGCGCGCGCGCGCSNNNDN	105169
OY		974	CGAGGGCGAGTGGACAATCCAATGGGCGTGGGCTTGCGCGCGCGCGTGTCTTTCGAC	1033
Db		105170	BGGGCGCGCGCGCGCGCSNNNDNCCBGGGCGCGCGCGCGCGCGCGCSNNNDNG	105229
OY		1034	CACCGCGGTGTATGAGCGCGCGCTGCGCAAGGCTATGACTGTTACAGGGCGACGC	1093
Db		105230	GCGCGCGCGCGCGCGCSNNNDNCCGCGCGCGCGCGCGCGCGCSNNNDNCCGCGCC	105289
OY		1094	GCTGCGCGCCTTGAGCTGAGCGAGGTGCGCGCATGACGACATCAAGAGACCAACAT	1153

Db	105280	GCCGGGCGGGCCGGSNNNDNNCGCCGGCCGGGCCGGSNNNDNNCGCCGGCCGGGGG	105349
Qy	1154	CACCATGCTGTCGGCGCCCGGGTCCAGTTTGAGAGAGCTGTGAAGCACGCGCGCGC	1213
Db	105350	CGSNNNDNNCGCGCGCGCCGGGCGGSNNNDNNCGCGCGCGCGGSNNNDNNCGCGG	105409
Qy	1214	GCGCGCGC-----AGGCGCGCGCGCACCGCACCCCCCGGGCCGCTGGCCTTGGAGCGG	1267
Db	105410	CCGGGCGCCSNNNDNNCGCGCGCGCGGCCGSNNNDNNCGCGCGCGCGSNNNDNNCGGCGG	105465
Qy	1268	CGCGGCGCTTACACAGCGAGAGACGGGAGGGGCGGATCACTGCGGCTGTGGCGTGSCAA	1327
Db	105470	GGCGGSNNNDNNCGCGCGCGCGCGSNNNDNNCGGCGGBGGCTTCGCTBSNNNDNNGGGCCCT	105529
Qy	1328	CGGCGTGGGCAACGCGCAAGAACTGATCACCAAGATGCAAGCGCGGCGAGCGCAATTACGA	1387
Db	105530	GGCTTGCGCCCCCGGGCCGGCTTGCCGCGCCGCGCGSNNNDNNGGGCCCTGAGCTCG	105589
Qy	1388	CTTTGTGAGATCATGCGCTGCCCCCGCGGCGTGTGGCGGCGGCGGCGCACGCCCCGCTC	1447
Db	105590	CCCGCGGCGCGGCTTGCGCGCGCGCGCGSNNNDNNGGGCGCCTTGACTCGGCGCGCGG	105649
Qy	1448	CACCG 1452	
Db	105650	CCCGG 105654	

RESULT 5.
 AAH13714
 ID AAH13714 standard; cDNA; 1562 BP.
 XX
 AC AAH13714;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:10601.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241699.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Oca T, Isogai T, Nishikawa T, Hayaeshi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 10601; 2537bp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

FT	/tag= a	
FT	/product= "Nysd2 partial protein"	
FT	/note= "CDS does not include stop codon"	
FT	complement (1056..2576)	
FT	/tag= b	
FT	/product= "Nysd1 protein"	
FT	2806..6906	
FT	/tag= c	
FT	/product= "Nysa protein"	
FT	6952..16530	
FT	/tag= d	
FT	/product= "Nysb protein"	
FT	16550..49840	
FT	/tag= e	
FT	/product= "Nysc protein"	
FT	50260..51015	
FT	/tag= f	
FT	/product= "Nyse protein"	
FT	51405..54305	
FT	/tag= g	
FT	/product= "Nysr1 protein"	
FT	54329..57190	
FT	/tag= h	
FT	/product= "Nysr2 protein"	
FT	/note= "CDS does not include start codon"	
FT	57180..59963	
FT	/tag= i	
FT	/product= "Nysr3 protein"	
FT	60415..61047	
FT	/tag= j	
FT	/product= "Nysr4 (short) protein"	
FT	/note= "CDS does not include start codon"	
FT	61736..62497	
FT	/tag= k	
FT	/product= "Nysr5 protein"	
FT	/note= "CDS does not include start codon"	
FT	complement (62551..63615)	
FT	/tag= l	
FT	/product= "Orp2 protein"	
FT	/note= "CDS does not include start codon"	
FT	63765..64961	
FT	/tag= m	
FT	/product= "Orp1 protein"	
PN	WO200159126-A2.	
PD	16-AUG-2001.	
XX		
PF	08-FEB-2001; 2001WO-GB00509.	
XX		
PR	08-FEB-2000; 2000GB-0002840.	
PR	10-APR-2000; 2000GB-0008786.	
PR	14-APR-2000; 2000GB-0009387.	
XX		
XX	(UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.	
PA	(SNTF) SINTER STIFTELSEN IND TEK FORSK.	
PA	(ALPH-) ALPHARMA AS.	
PA	(SINV-) SINVENT AS.	
PA	(DZIE/) DZIELEWSKA H.	
PA	(ZOTC/) ZOTCHEV S B.	
PA	(SEKU/) SEKUROVA O N.	
PA	(FJAE/) FJAEVRIK E.	
PA	(BRAU/) BRAUTASET T.	
PA	(STRO/) STROM A R.	
XX		
PI	Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR.	
XX	Valla S, Ellingsen TE, Sletta H, Gulliksen O;	
XX	WPI: 2001-557614/62.	
DR	P-PBDB: AAE10125, AAE10126, AAE10127, AAE10128, AAE10129, AAE10130,	
DR	AAE10131, AAE10132, AAE10133, AAE10134, AAE10135, AAE10136, AAE10137	
PT	New nystatin polyketide synthase polynucleotides and polypeptides,	

PT	useful as antibiotics and antifungals -
XX	
PS	Claim 2; Page 116-151; 266pp; English.
XX	
CC	The present invention relates to the cloning and sequencing of the gene
CC	cluster encoding a modular type I polyketide synthase (PKS) enzyme
CC	involved in the biosynthesis of the macrocyclic antibiotic nystatin.
CC	The nystatin PKS is useful as antifungal antibiotics. The present
CC	sequence is a Streptomyces noursei nysI DNA of nystacin PKS gene cluster.
XX	
SQ	Sequence 65140 BP; 8270 A; 25171 C; 22273 G; 9426 T; 0 other;
Query March	2.9%; Score NO. 2; DB 22; Length 65140;
Best Local Similarity	42.5%; Pred. No. 0.0015;
Matches 573; Conservative	0; Mismatches 758; Indels 18; Gaps 3;
OY	226 AGGTCGCCCGCCCCGGCTCCGCTGCAGCAGACCGCGTGTGTAGCCCTTGCAACTTG 285
DB	18456 AGCTTCGCGCGCGGCGACATGTCGCGGGGTCTCTCCTCGAACAAGCTTGACCCCTGCG 18515
OY	286 AGGCGCCCGGACGCGCGCCTTAGGCAAGTGCCTTGCGCGGCTGCCAACCGGCTGCGAGG 345
DB	18516 CCGGCGCGCGCGCGCTCATGACAGCCCTGCGCGGCGCGCGCGATGCTCGGATCCGCG 18575
OY	346 CGCCTTTGAATCATGTTCACAGAGCGCGCTGCGCGGATTGCCAAAGCCCAAGATCAACCCA 405
DB	18576 CCACCGAGGACGAGGTGATCCCGCCCACTTACCGACGACTCTTGATTCGCGCGCTTCAA 18635
OY	406 CGGCGAAGGACGATCTGTGGTGCAGATGGACTCCGGCCGTTGTGCTATTGCTCCAGAACCC 465
DB	18636 GGCCCACTTCCTGTCGTGTGTGCGGACCGAAGGAACCTGCGCGGATCGGGGGCGGCT 18695
OY	466 TGCGGCTTGCGCGCGCGGCGCCACCACCCCAGAAGCTGGCCGAGGAGCTTCGCGCGCTCG 525
DB	18696 TCACCGGCCAGAGACCCGAAAGACCACCGCGCTGGGGGTGACGACCGCTTCCACTGCGCG 18755
OY	526 GCCTTGAACGAGGTGTTG-----AACGCTGTTTGGCGCGCGCACTGACATCATGTGAGG 579
DB	18756 TCATGAGCCCGGATGCTGTGGCGGAATTCGCGCGGTGCTGCGCGGCGCTGATCTACAGAGC 18815
OY	580 AGGCGACAGAGCTGTGCACCGGCTCACCGAGACCTGTGAGGCGCCACCGGCACTCCGACG 639
DB	18816 CGGCGATCCCGGTCTCTTCCAACTTACCGGACCGTCCGCGCGTGGCGGACCTGTGCT 18875
OY	640 AGCGGCTGCCCATGTTTACCAAGCTGTGC---CCGGCTGATCGCTATGCTGAGAAAT 696
DB	18876 CCGCGCATGTAGGGGTCCGCACTGCCGAGGCGGCTCGCTTGGCGGACGGCGTCAACG 18935
OY	697 CTTAACCGGACCTGATTCCTCTTACGTAGACGTGCAGAGCGCCCGCAGATGATGCTGGCG 756
DB	18936 CCTTCACCGACCGCGCGGTGATCACGCTGTCGAACTCGGCGCGGCGGCTGTGTCCG 18995
OY	757 CCATGTGCAAGTGCCTTACTGAGGGAAGAAGGCGATTCGCGCAAGACATGTGTCATCG 816
DB	18996 CCAATGGCCCAAGAAATCTCTTCCGAGCGGCGCGCGCGCTGTGCGCTGTGCGCAAGACC 19055
OY	817 TGTTCATCATGACCTTGCACGCGGCAAGCAAGTCGAGGCTGACCGGACTGGTTCGTGTGG 876
DB	19056 GCCCGAGAGAGCTTCGCGCGGTACCGGCGCTGGCGCGCGCCACGTCGCGGCGTACGG 19115
OY	877 AGCGCGAACCCCACTTCCTGCGCGCACTGGACAGCTATCATCACACCGTGTGAGAGCTGGCA 936
DB	19116 TCCGCTGAGGCGCGCTTCTTGAAGGACACCGGCGCGCGCGCGGCACTGTGCCACTTAC 19175
OY	937 TCTTCAGAGAGCGCGGATCAACTGAGCGAGTCGCGAGGCGAGGTGGCAATTCMA 996
DB	19176 CTTTCAGACACAGGATTTTGGCCAGCGCGGCGCGCGCGCTCCAGACGTCAACGCGG 19235
OY	997 TGAGCGTGTGGCTTCGAGCGCGCGGCTGTTCGGAACAACCGAGCGGTGTCAATGAGAGCGG 1056
DB	19236 CGGGAATGTGGGCGCGCGGACCAACCGGCTGTGGGCGCAACGTTCGAATCTCGCGACGGG 19299
OY	1057 CGTTCGCGACGCGCTTATAGCTTTACGAGGACGCGCGCTGCGCGCTTGAAGCTTGAAGC 1116


```

Db      19296 CCGGCTACTTGTTCACACAGCGGCTCTCGTCCGGACCCACCTTGCTGCGGACACG 19355
OY      1117 AGGTGCGCGCATGACGAGCATCAAGAGACCAACATCAACATGATGCGCGCGCGG-- 1174
Db      19356 GGGTCCAGGGCGGGGCGCTGCTGCTGCGGACCCCTTCTGCAACTGAGCTGCTGCGCGCG 19415
OY      1175 -----GTCCAAATTGAGAGAGCTGCTGAAGACCGCGCGCGCGCGCGCGCGCGCG 1227
Db      19416 GCGACGAGCGCGCTGCGGACCGCGTCAAGAGAACTGACCTTGCGCGCGCGCGCGCG 19475
OY      1228 CCGCGACGCGACCGCGCGCGCGCGCGCTGCGCTGCGAGACCGCGCGCGCGCTTACCAAGCAGG 1287
Db      19476 CCGAGCGCGCGCGCTGCTCAACTCCAGGTCCGCGCTGCGCGCGCGCGCGCGCGCG 19535
OY      1288 AGCGAGCGCGCGCATCACACTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1347
Db      19536 GCACTCTCGGCACTTCTTCCCGGTGAGAGACGCTTCCGACCTGCTGCTGCGCAACAG 19595
OY      1348 AGCTGATCAACAAATGACAGCGCGCGCGCGCGCAAGTACGACTTGTGAGATGATGCGCT 1407
Db      19596 CCACCGCGCTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 19655
OY      1408 GCGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1467
Db      19656 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19715
OY      1468 AGAAGCGCGCGCGCGCTGTACACCTGAGAGAGAGTCCAGCGCTGCGCGCGCGCG 1527
Db      19716 TCGGCTTCCAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19775
OY      1528 AGAACCGTCCATCCGCGAGCTGTAGAC 1556
Db      19776 AGGTTACCGCGGAAGTGGCGCTGCCGAC 19804

RESULT 8
AAD17186
ID      AAD17186 standard; DNA; 125401 BP.
XX
XX      AAD17186;
AC
XX
XX      29-NOV-2001 (first entry)
DT
XX
XX      Streptomycetes noursei nystatin PKS gene cluster DNA.
DE
XX
XX      Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;
KM      antifungal; antibiotic; ds.
XX
XX      Streptomycetes noursei.
OS
XX
XX      Key
FH      Location/Qualifiers
FT      CDS
FT      6337..34771
FT      /tag= a
FT      /product= "Nysi complete protein"
FT      34792..51099
FT      /tag= b
FT      /product= "NysJ protein"
FT      51155..57355
FT      /tag= c
FT      /product= "NysK protein"
FT      57503..58687
FT      /tag= d
FT      /product= "NysL protein"
FT      complement (58786..58980)
FT      /tag= e
FT      /product= "NysM protein"
FT      /note= "CDS does not include start codon"
FT      complement (59045..60241)
FT      /tag= f
FT      /product= "NysN protein"
FT      /note= "CDS does not include start codon"
FT      complement (60238..61296)
FT      CDS

```

```

FT      /tag= g
FT      /product= "NysD2 complete protein"
FT      120628..121308
FT      /tag= h
FT      /product= "NysR4 (long) protein"
XX
XX      W0200159126-A2.
XX
XX      16-AUG-2001.
XX
XX      08-FEB-2001; 2001WO-GB00509.
XX
XX      08-FEB-2000; 2000GB-0002840.
XX      10-APR-2000; 2000GB-0008786.
XX      14-APR-2000; 2000GB-0009387.
XX
XX      (UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
XX      (SNIF) SINTER STIFTELSEN IND TEK FORSK.
XX      (ALPH-) ALPHARMA AS.
XX      (SINV-) SINVENT AS.
XX      (DZIE/) DZIELEWSKA H.
XX      (ZOTC/) ZOTCHEV S B.
XX      (SEKU/) SEKUROVA O N.
XX      (FJAE/) FJAEVRIK E.
XX      (BRAU/) BRAUTASET T.
XX      (STRO/) STROM A R.
XX
XX      Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
XX      Valla S, Ellingsen TE, Sletta H, Gulliksen O;
XX
XX      WPI: 2001-557614/62.
XX      P-PSDB; AAE10143, AAE10144, AAE10145, AAE10146, AAE10147, AAE10148,
XX      AAE10149, AAE10150.
XX
XX      New nystatin polyketide synthase polynucleotides and polypeptides,
XX      useful as antibiotics and antifungals -
XX
XX      Claim 1; Page 188-254; 266p; English.
XX
XX      The present invention relates to the cloning and sequencing of the gene
XX      cluster encoding a modular type I polyketide synthase (PKS) enzyme
XX      involved in the biosynthesis of the macrolide antibiotic nystatin.
XX      The nystatin PKS is useful as antifungal antibiotic. The present
XX      sequence is a Streptomycetes noursei nystatin PKS gene cluster DNA.
XX
XX      Sequence 125401 BP; 15664 A; 49692 C; 42871 G; 17174 T; 0 other;
XX
XX      Query Match
XX      Best Local Similarity 2.9%; Score 70.2; DB 22; Length 125401;
XX      Matches 573; Conservative 0; Mismatches 758; Indels 18; Gaps 3;
OY      226 AGGTGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 285
Db      78717 AGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 78776
OY      286 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 345
Db      78777 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 78836
OY      346 CGCTTTGAGTATGTCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 405
Db      78837 CCACCGAGAGCGAGTACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 78896
OY      406 CCGCGAGACGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 465
Db      78897 GGGCCACCTTCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 78956
OY      466 TGGGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 525
Db      78957 TCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 79016
OY      526 GCTTTAGCAGGCTGTTG-----ACAGCGTGTGCGCGCGCGCGCGCGCGCGCGCG 579

```

```

Db 79017 TCATGAGCCGATGCTGCGGAATTCGCCGCCCTGCCGCCGCGCTGACTTACCAAGAC 79076
QY 580 AGGCGACGAGCTGCTGCACACCGCTTACCGAGCACTTGAGGCGCACTCCGACTCCGACG 639
Db 79077 CGGCAATCCCGGCTCTCTCAACCTCACGCGCACCGCGCGCGCTGCGGACTGTGCT 79136
QY 640 AGCGCGTCCCATGTTTACAGAGTGTGC---CCGCGTGTGATGCTGATGCTGGAATAAT 696
Db 79137 CGGCGCACTTACTGGGTCCGCAAGTCCGCGAGGCGGTGCTGCGCAAGCGGTACCG 79196
QY 697 CTACCCCGGAGCTGATCTCCCTTACGTAGAGCTGCAAGAGCCCGCAATGATGCTGCGCG 756
Db 79197 CCTTACACGACCGCGCGGTGACACCGTGTGCAATCTCGCGCGGAGCGGTGTCTGCG 79256
QY 757 CCAATGTCAGATCTTACTTACGCGAAAAGAGGCAATCGCGCAAAAGACATGCTATG 816
Db 79257 CCAATGCGCCAGAAATCCCTGCGGAGCGCGCGCGCGCGCTGCGCTGCTGCGCAAGAC 79316
QY 817 TGTTCATCATGCTCTGACGCGGCAAGCAAGTCCGAGAGCTGACCGCGACTGTTCTGTG 876
Db 79317 GCCCGAGAGAGCTCTCGCGCTGACCGGCTGCGCGCGCGCGCTGCGCGGTGACG 79376
QY 877 AGCGGACCCGACCTGCGCGAGCTGGAACCACTCATCAACCGTGAAGCTGAGCAACA 936
Db 79377 TCCGCTGCGCGCTGCTTTCAGAGGACCGCGCGCGCGCGCGCTGCGCGCTGCG 79436
QY 937 TCTTCAAGAGCGCGGATACCTGCGCGAGCTGCGCGAGGCGAGTGGACATTCGA 996
Db 79437 CTTTCCAGACACAGGCTTTCGCGCGACCGCGCGCGCGCGCGCGCGCGAGATCACGCG 79496
QY 997 TGGCGTGGGCTGCGCGCGCGCGCGCTGTGTTCCGACACCGCGCGGTGATGAGAGCG 1056
Db 79497 CGGAGCTGGCGCGCGCGCGACCGCTGCTGCGCGCACTGCGAGCGAGG 79556
QY 1057 CGCTGCGCACGCGCTTATGAGCTGTTTCAAGGCGACGCGCGCGCGCGCTGAGCTGAGCG 1116
Db 79557 CCGGCTACTTGTTCACAGCGCGCTTCCGTCCGACCCACCTGCTGCGCGACGAC 79616
QY 1117 AGGTGCGCGCATGAGCGGATCAGAGAGACCAATCATGATGTCGCGCGCGCG-- 1174
Db 79617 GGGTTCAGGCGCGCGCGCTGCTGCGCGCGACCGCTTGTGAACTGCGCGCTGCGCGCG 79676
QY 1175 -----GTCCAAATTGAGAGAGCTGTGAAAGACCGCGCGCGCGCGCGCGCGCGCG 1227
Db 79677 GCGACGAGCGCGCTGCGACCGGTGAGGAACTGAGCTTGGCGCGCGCGCTGCTGCTG 79736
QY 1228 CCGCGCACGCGACCGCGCGCGCGCTGAGCTGAGAGCGCGCGCGCGCTTACCAAGAG 1287
Db 79737 CCGAGCGCGCGCGCGCTCAACTCAAGTCCGCGTCCGCGCGCGCGCGCGCGCGCGCG 79796
QY 1288 ACGGCGAGGCGCGCATCAGCTGCGGTGCGCTGAGCGCAAGGCGTGGGCAAGCCCAAG 1347
Db 79797 GCACCTTCCGCGCATCTTCTCCGCGGTGAGAGACGCGCTTCAAGCTGCGCTGCTCA 79856
QY 1348 AGCTGATCAACAGATCAGAGCGCGCGCGCAAGGCGCAAGTTCGTTGGAATATGCGCT 1407
Db 79857 CCAACCGCGCTGTAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 79916
QY 1408 GCGCGCGCGCGCTGTTGAGGCGCGCGCGCGCGCGCGCGCGCGCTTCAAGCGCAATCAGC 1467
Db 79917 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 79976
QY 1468 AGAAGCGGAGGCGCGCGCTGTAACAAGTGAAGAAAGTCAAGCTGCGCGCGCGCGCG 1527
Db 79977 TCGGCTTCAAGTACGCGCGCGCGCTTTCAGAGGCGCTGCGCGCGCGCGCGCGCG 80036
QY 1528 AGAAGCGGCTCATCGCGAGCTGACAGC 1556
Db 80037 AGGTCTAGCGCGAGAGTGGCGCTGCGCGAC 80065

```

RESULT 9
AAAT0438

```

ID AAA70438 standard; DNA; 1836 BP.
XX
AC AAA70438;
XX
DT 28-NOV-2000 (first entry)
XX
DE HIV gp140.mut.mod5F162.delV1/V2 construct coding sequence.
XX
KW HIV-1; AIDS; Gag; vaccine; expression cassette; Env; ss.
XX
OS Human immunodeficiency virus type 1.
XX
PN MO200039302-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WC-US31245.
XX
PR 31-DEC-1998; 98US-0114495.
XX
PR 01-DEC-1999; 99US-0168471.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;
PI Greer C, Selby M, Walker C;
XX
DR MPI; 2000-452400/39.
XX
PT Expression cassettes encoding the human immunodeficiency virus (HIV)
PT Gag-containing polypeptide useful for vaccinating against HIV
PT infections and acquired immunodeficiency syndrome (AIDS) -
XX
PS Claim 22; Fig 28; 391pp; English.
XX
CC The present invention relates to synthetic HIV protein expression
CC cassettes. The present sequence is the coding sequence of a synthetic
CC HIV Env expression cassette. In addition, synthetic HIV Gag expression
CC cassettes have been constructed. The Gag protein of HIV is needed for
CC the assembly of virus-like particles. In addition, the Gag protein is
CC involved in many stages of the HIV life cycle, including assembly, virion
CC maturation after particle release and early post-entry steps in viral
CC replication. The expression cassettes of the present invention may be
CC used for the recombinant expression of HIV Gag-polypeptides which may
CC then be used to vaccinate against HIV infection and acquired
CC immunodeficiency syndrome (AIDS).
XX
SQ Sequence 1836 BP; 414 A; 623 C; 550 G; 249 T; 0 other;

```

Query Match 2.9%; Score 69.8; DB 21; Length 1836;
Best Local Similarity 44.3%; Pred. No. 0.0015;
Matches 422; Conservative 0; Mismatches 522; Indels 9; Gaps 3;

```

QY 462 ACCCTGGGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521
Db 544 ACCCAGGCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 603
QY 522 CTCGCTTTCAGAGAGTGTTCAGACGCTGTTTGGCGCGCGCGCGCGCGCGCGCGCG 581
Db 604 GAGGCGGTGGTATTCGCGAGCGAGGAACTTACCGAGCAAGCGCAAGCAACATCATGTCAG 663
QY 582 GCGAGGAGAGTGTTCAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 641
Db 664 CTGAAGAGAGCGGTGAGATCAACTGACCTCCCGCAACAAACCGCGAGAGATC 723
QY 642 CCGTCCCATGTTTCAACAAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 701
Db 724 ACCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 783
QY 702 CCGGACTGATTCCTTACGTGAGAGCTGCAAGAGCGCGCGCGCGCGCGCGCGCGCG 761
Db 784 GCCCATGCAACATCAAGCGCGCGAGAGTGGAAACAAACCTTGAAGAGATCGTACCAAG 843

```


Db 1312 ATGCGGACAACTGGCGGACGAGCTGTACAGTACAGGTGTGAAGTCCGCTTG 1371
QY 1176 TCCAGATTGAGAGCTGTGAAGCAACCGCGCGCGCGCGCGCGCGCGCGCGC 1235
Db 1372 GGGGTGGCCCCCAAGAGCGCAAGCGCGGTGTGACAGCGAAGCGCGCGTGC 1431
QY 1236 GGGACCCCCCGCGCGCTGTGGAGCGCGCGCGCGCTTCAACGAGAGCGCGAG 1295
Db 1432 CTGGGCGCCATGTCTTCTGGGCTTCTGGGCGCGCGCGCGCGCGCGCGCG 1491
QY 1296 GGGGCGATCATCTGGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1355
Db 1492 CTGACCTGACCGTGA---GGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1548
QY 1356 ACCAAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1408
Db 1549 AACCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1601

RESULT 11
AAA70435
ID AAA70435 standard; DNA; 1944 BP.
XX AAA70435;
AC
XX
DT 28-NOV-2000 (first entry)
XX
XX HIV gp140.modsF162.delV1/V2 construct coding sequence.
XX
XX HIV-1; AIDS; Gag; vaccine; expression cassette; Env; SB.
XX
OS Human immunodeficiency virus type 1.
XX
XX Synthetic.
XX
PN W0200039302-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US1245.
XX
PR 31-DEC-1998; 98US-0114495.
XX 01-DEC-1999; 99US-0168471.
XX
PA (CHIR), CHIRON CORP.
XX
PI Barnett S, Zur Megede J, Srivastava I, Lian Y, Harzog K, Liu H;
PI Greer C, Selby M, Walker C;
XX
DR WPI; 2000-452400/39.
XX
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
PT Gag-containing polypeptide useful for vaccinating against HIV
PT infections and acquired immunodeficiency syndrome (AIDS) -
XX
PS Claim 22; Fig 25; 391p; English.
XX
XX The present invention relates to synthetic HIV protein expression
CC cassettes. The present sequence is the coding sequence of a synthetic
CC HIV Env expression cassette. In addition, synthetic HIV Gag expression
CC cassettes have been constructed. The Gag protein of HIV is needed for
CC the assembly of virus-like particles. In addition, the Gag protein is
CC involved in many stages of the HIV life cycle, including assembly, virion
CC maturation after particle release and early post-entry steps in viral
CC replication. The expression cassettes of the present invention may be
CC used for the recombinant expression of HIV Gag-polypeptides which may
CC then be used to vaccinate against HIV infection and acquired
CC immunodeficiency syndrome (AIDS).
XX
SQ Sequence 1944 BP; 450 A; 656 C; 577 G; 261 T; 0 other;

Query Match 2.9%; Score 69.8; DB 21; Length 1944;
Best Local Similarity 44.3%; Pred. No. 0.0015;
Matches 422; Conservative 0; Mismatches 522; Indels 9; Gaps 3;

QY 462 ACCCTGGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521
Db 652 ACCCAAGGATCG 711
QY 522 CTGGGCTTTGAGAGGTGTTTACAGCGGTGTTTGGGCGGACCTGACCATCATGAGAG 581
Db 712 GAGGGGTGTGTATCCGACGCGAGATTTACCGACAAAGCGCCAAACCATCATCTGTGCG 771
QY 582 GGCAGGAGCTGTGTACACCGCGCTTACCGAGCACTGGAGGCCACCGCATCTCGAGCG 641
Db 772 CTGAAGAGAGGTGAGATCAATGTGACACCGCGCGCGCGCGCGCGCGCGCGCG 831
QY 642 CCGCTGCCATTTTCAACAGCTGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 701
Db 832 ACCATGCG 891
QY 702 CCGGACCTGATCCCTTACGTGAGCAGCTGCAAGAGCGCGCGCGCGCGCGCGCG 761
Db 892 GCCCACTGCAATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 951
QY 762 GTCAAGTCTTACCTAGCGGAAAGAGGCAATCGCGCGCGCGCGCGCGCGCGCGCG 821
Db 952 CTGCAAGGCCCATGTGGGCAAGAGCAATCGTGTTCAGCAGAGCGCGCGCGCGCGCG 1011
QY 822 ATCATGCGCTGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 881
Db 1012 GAGATCGTGTATCAGACAGCTTCAATGCGCGCGCGCGCGCGCGCGCGCGCGCG 1071
QY 882 GACCCACCTCG 941
Db 1072 CTGTTTCAACAGCACTGGAACAAACATCGCGCGCGCGCGCGCGCGCGCGCGCG 1131
QY 942 AAGAGCGCGGATCAACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1001
Db 1132 CTGCG 1191
QY 1002 GTGGGCTCG 1058
Db 1192 GCCCG 1251
QY 1059 CTGGCGAGGCGCTATAGCGTGTTCAGCGGCGCGCGCGCGCGCGCGCGCGCGCG 1118
Db 1252 CGGACG 1311
QY 1119 GTGCGCGGATGAGCGGATCAAGAGCAACATCACCAGT---GCCCGCGCGCGCG 1175
Db 1312 ATGCGGACAACTGGCGGACGCGGTGTACAAAGTGAAGTGAATCGAGCGCGCTG 1371
QY 1176 TCCAGATTGAGAGCTGTGAAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1235
Db 1372 GGGGTGGCCCCCAAGAGCGCAAGCGCGGTGTGACAGCGAAGCGCGCGTGC 1431
QY 1236 GGCACCG 1295
Db 1432 CTGGGCGCCATGTCTTCTGGGCTTCTGGGCGCGCGCGCGCGCGCGCGCGCG 1491
QY 1296 GGGGCGATCATCTGGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1355
Db 1492 CTGACCTGACCGTGA---GGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1548
QY 1356 ACCAAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1408
Db 1549 AACCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1601

RESULT 12
AAA70437
ID AAA70437 standard; DNA; 1944 BP.
XX
XX AAA70437;
AC
XX
DT 28-NOV-2000 (first entry)

```

XX HIV gp140.mut.modsF162.delV2 construct coding sequence.
DE
XX HIV-1; AIDS; Gag; vaccine; expression cassette; Env; ss.
XX
XX Human immunodeficiency virus type 1.
OS
XX Synthetic.
XX
XX WO200039302-A2.
XX
XX 06-JUL-2000.
XX
XX 30-DEC-1999; 99WO-US31245.
XX
XX 31-DEC-1998; 98US-0114495.
XX
XX 01-DEC-1999; 99US-0168471.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;
PI Greer C, Selby M, Walker C;
PI
XX WPI; 2000-452400/39.
XX
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
PT Gag-containing polypeptide useful for vaccinating against HIV
PT infections and acquired immunodeficiency syndrome (AIDS) -
XX
XX Claim 19; Fig 27; 391pp; English.
XX
XX The present invention relates to synthetic HIV protein expression
XX cassettes. The present sequence is the coding sequence of a synthetic
XX HIV Env expression cassette. In addition, synthetic HIV Gag expression
XX cassettes have been constructed. The Gag protein of HIV is needed for
XX the assembly of virus-like particles. In addition, the Gag protein is
XX involved in many stages of the HIV life cycle, including assembly, virion
XX maturation after particle release and early post-entry steps in viral
XX replication. The expression cassettes of the present invention may be
XX used for the recombinant expression of HIV Gag-polypeptides which may
XX immunodeficiency syndrome (AIDS).
XX
XX Sequence 1944 BP; 451 A; 655 C; 577 G; 261 T; 0 other;
SQ
Query Match 2.9%; Score 69.8; DB 21; Length 1944;
Best Local Similarity 44.3%; Pred. No. 0.0015;
Matches 422; Conservative 0; Mismatches 522; Indels 9; Gaps 3;
OY 462 ACCCTGGGCTGTGGCGCGCGCCACCACCCCAAGCAGCTGGCGAGGCTCCGCGCG 521
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 652 ACCCAAGGATCCGCCCGGTGAGACACCACTGCTGTGAACGGCAGCTGGCGAG 711
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 522 CTCGGCTTTGACGAGGTGTTTACACGCTTTTGGCGCGCAGCTGACCATGAGAG 561
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 712 GAGGGGTGTGTATCCGACGAGAACTTCAACGAGAACCCCAAGCATCTGTCAG 771
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 582 GCGAGGAGCTGTGACCGGCTCACGAGCAGCTGAGGCGCCACCGGACCTCCGAG 641
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 772 CTGAGGAGAGCTGAGATCACTGACCCGCCCAACAACACCCGCAAGAGATC 821
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 642 CCGCTGCCATGTTCAACGAGCTGTGCCCGGAGTGTGCTGTGAGAAATCTTAC 701
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 832 ACCATGCGCCCGCGCGCGCTTCTAGCCACCGGAGCATCTAGGCGCATCCCGCAG 891
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 702 CCGGAGCTGATCCCTCACTGAGCAGCTGCAAGAGCCCAATGATGCTGGCGCATG 761
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 892 GCCCATGCAACATCAAGCGCGGAGAGTGAACAACCTTGAAGAGATCTGACCAAG 951
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 762 GTCAAGCTTACCTAGCGGAAAGAGGAGCATGGCGCAAGAGCATGTGATGGTGC 821
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 952 CTGAGGCGCATGTTGCGCAACAGACCATGTGTTCAAGAGAGAGCGCGCGACCCC 1011
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 822 ATCATGCTGTGACGGCGCAAGCATGAGGCTGACCGGACATGTCTGTGTGAGCGCC 881

```

```

DB 1012 GAGATGTGATCACAAGCTTCAACTGGGCGCGAGTTTCTTACTGCAACAGCACCCAG 1071
OY 882 GACCCCACTTGGCGAGCTGAGACCAAGTCATACCAAGCTGGAGCTGGCAACATCTTC 941
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1072 CTGTTCAACAGCACTGGAAACAACACCATGAGCCCAACACCAAGCAGCATCTACC 1131
OY 942 AAGAGCGCGGATCACTGAGCGAGCTGCGGAGGCGAGTGGGACAATCCAAATGGG 1001
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1132 CTGCGCTGGCATTAAGAGATCAACCGCTGCAAGAGTGGGCAAGGCATGTAC 1191
OY 1002 GTGGGCTGGGCGCGCG--GTGCTGTTCGCAACCGCGGTGTCTATGAGCGCGG 1058
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1192 GCCCGCCCATCGCGCGAGATCCGCTGACAGCAACATCACCGGCTGTGTGAC 1251
OY 1059 CTGCGAGGCGCTATGAGCTGTTCAGGAGCAGCGCTCGCGCTGAGCGTGAAGGAG 1118
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1252 CCGACGCGCGCAAGAGATCAACACCAACCAAGATTTCTCGCCCGCGCGGCGAGC 1311
OY 1119 GTGCGGAGTGAAGCGCATCAAGAGAACCAATCACCATGT--GCCCGCGCGGAG 1175
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1312 ATGCGGACACATGGGCGAGCGAGCTGTACAAAGTACAAAGTGAAGATCGAGCCCTG 1371
OY 1176 TCCAGTTTGAAGAGCTGCTGAAGACCGCGCGCGCGCGCGCGCGCGCGCGAC 1235
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1372 GCGGTGGCCCCACCAAGGCCAAGCGCCCGGTGTGACCGGAGAAAGCGCGGTGACC 1431
OY 1236 GCGACCCCGCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1295
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1432 CTGGGTGCATGTTCTTGAGCTTCTGCGCGCGCGCGCGAGACATGAGCGCCGAGC 1491
OY 1296 GCGGAGTCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1355
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1492 CTGACCTGACGTGA---GCGCGCGAGCTGTGAGGCGATCTGCGAGCGAGAAC 1548
OY 1356 ACCAAGATCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1408
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1549 AACCTGCTGCGGCGCATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1601

```

RESULT 13
AAA70433
ID .AAA70433 standard; DNA; 2025 BP.
XX
XX AAA70433;
AC
XX
DT 28-NOV-2000 (first entry)
XX
XX HIV gp140.modsF162 construct coding sequence.
DE
XX HIV-1; AIDS; Gag; vaccine; expression cassette; Env; ss.
XX
XX Human immunodeficiency virus type 1.
OS
XX Synthetic.
XX
XX WO200039302-A2.
XX
XX 06-JUL-2000.
XX
XX 30-DEC-1999; 99WO-US31245.
XX
XX 31-DEC-1998; 98US-0114495.
XX
XX 01-DEC-1999; 99US-0168471.
XX
XX (CHIR) CHIRON CORP.
XX
XX Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;
PI Greer C, Selby M, Walker C;
PI
XX WPI; 2000-452400/39.
XX
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
PT Gag-containing polypeptide useful for vaccinating against HIV


```

OY 522 CTCGGCTTTGACGAGGTGTTTGAACAGCTGTTTGGGCGGACCTGACATCATGAGAG 581
DB 793 GAGGCGTGTGTATCCGACGAGAACTTACCGAACAAGCCCAACATCATCTGTCAG 852
OY 582 GGCAGGAGCTGTGTGACCGGCTTACCGAGACCTGTGAGGCCCAACCCGATCTCCGACGAG 641
DB 853 CTGAAGAGAGAGCTGTGAGATCAACCTGACCCCGCCCAACAACACCCGCAAGAGCATTC 912
OY 642 CCGGCTCCCATGTTTCAACAGCTGTGCCCCGGGTGTGATGCTTGTGAGAAATCTTAC 701
DB 913 ACCATGGGCCCCGGCGCGCTTCTAGCGACCGGAGCATCTACGGGACATCCGCCAG 972
OY 702 CCGGACCTGATCCCTTACCTGAGCAGCTGCAAGAGCCCGCATGTACTGGCGGCATG 761
DB 973 GCCCACTGCAACATCAGCGCGGAGAAAGTGAGAACACCTCTGAAGAGATCTGACCAAG 1032
OY 762 GTCAACTCTTACCTAGCGGAAAGAGGATGGCCCAAGACATGTGATGTGTC 821
DB 1033 CTGCAAGGCCAGTTGGCAACAAGACCATGTGTTCAAGCGAGAGCGGCGGACCC 1092
OY 822 ATCATGCTGTGACGCGCAAGCAGTGTGAGGCTGACCGGACATGTTCTGTGTGACGCGC 881
DB 1093 GAGATGTGTATGACAGCTTCAACTGCGGCGGAGATTTCTTACTGTCAACAGACCCAG 1152
OY 882 GACCCCACTGTGCGCAGCTGTGACCAACGTCATCACCAGCTGTGAGCTGGCAATCTTC 941
DB 1153 CTGTTCACACAGCAGCTGTGAACAACACCATGTGCGCCCAACAACACGACCATCATCAC 1212
OY 942 AAGGAGCGGCGCATCAACCTGTGCGGAGCTGCGGAGGCGAGTGTGAGCAATCCAAATGGGC 1001
DB 1213 CTGCGCTGTGCGCATCAAGAGATCATCAACCGTGTGAGAGGTGTGGCAAGGCGCATGTAC 1272
OY 1002 GTGGGCTGTGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1058
DB 1273 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1332
OY 1059 CTGCGGACGCGCTGTAGCTGTTCACGCGGACCGCGCTGCGCGCGCTGTAGCTGTAGCGAG 1118
DB 1333 CGGACGCGGCGGCAAGAGATCAACCAACCGAGATCTTCGCGCGCGCGCGCGCGCGCGCGG 1392
OY 1119 GTGGGCGGCGCATGAGCGGCGCATCAAGAGAGACCAATCAATCAATGAT---GCGCGCGCGCGG 1175
DB 1393 ATGCGGAGCAACTGTGGCGGAGCGGCTGTACAACTGTGTGAAGATCGAGCCCTG 1452
OY 1176 TCCAAATTGAGAGAGCTGTGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1235
DB 1453 GCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1512
OY 1236 GGCACCGCGCGCGCGCGCTGTGCGGAGCGGCGCGCGCGCTTCAACGAGCGAGCGAGG 1295
DB 1513 CTGGGCGCGCATGTTCCTGGGCTTCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1572
OY 1296 GCGGCGCATCACTGTGCGCGGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1355
DB 1573 CTGACCTGTACCTGTGA---GGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1629
OY 1356 ACCAAGATGAGCGCGCGCGCGAGCAAGTACAGCTTTGTGAGATCAATGCGCTG 1408
DB 1630 AACCTGTGTGCGCGCATGTAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1682

```

RESULT 15
 AAAS1944 standard; DNA; 2310 BP.

AAAS1944;
 31-OCT-2000 (first entry)
 Modified HIV-1 Env gp160, Val120-A1a204.
 HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env;
 CD4 binding region; VI/V2 loop; bridging sheet; ss.

```

XX OS Human immunodeficiency virus type 1 isolate SF162.
OS Synthetic.
XX PN MO200039303-A2.
XX PD 06-JUL-2000.
XX PF 30-DEC-1999; 99WO-US31272.
XX PR 31-DEC-1998; 98US-0114495.
XX PR 29-SEP-1999; 99US-0156670.
XX PA (CHIR ) CHIRON CORP.
XX PI Barnett S, Hartog K, Martin E,
XX WPI, 2000-465745/40.
XX DR
XX PT Novel modified HIV Env polypeptides useful as immunizing agents and for
XX PT preparing a vaccine to elicit an immune response against a broad range
XX PT of HIV subtypes
XX PS Claim 14, Page 118; 139pp; English.
XX PS
XX CC Novel immunogenic modified human immunodeficiency virus (HIV) envelope
XX CC (Env) polypeptides having an amino acid deleted or replaced in the region
XX CC corresponding to residues 420-436 or 119-123 and 199-210 relative to
XX CC isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV
XX CC strain SF162, with numbering relative to isolate HXB-2. The Env
XX CC polypeptides are modified so as to expose at least part of the CD4
XX CC binding region. The modified HIV Env polypeptides, coding polynucleotides
XX CC and constructs, further comprising an adjuvant, are used for inducing an
XX CC immune response in an individual. The method involves administering a
XX CC first composition comprising a polynucleotide encoding the Env
XX CC polypeptide in a priming step and administering a second composition
XX CC comprising a modified Env polypeptide as a booster in an amount
XX CC sufficient to induce an immune response in the individual. The first
XX CC and/or second composition further comprises an adjuvant (claimed). The
XX CC intracellularly produced Env polypeptides can be used for a number of
XX CC diagnostic and therapeutic purposes to determine the presence of reactive
XX CC antibodies and/or Env proteins in a biological sample to aid in the
XX CC diagnosis of HIV infection or disease status or as measure of response to
XX CC immunization.
XX XX
XX Sequence 2310 BP; 477 A; 805 C; 707 G; 321 T; 0 other:
XX
XX Query Match 2.9%; Score 69.8; DB 21; Length 2310;
XX Best Local Similarity 44.3%; Pred. No. 0.0015;
XX Matches 422; Conservative 0; Mismatches 522; Indels 9; Gaps 3;

```

```
Db      796 CTGCAGGCCCAAGTTCCGCAACAAGACCATCGTGTTCAGCAGACGACGGCGGACCCC 855
QY      822 ATCATGCCCTGCACGCGCAAGCAGTCGAGGCTGACCGGCACTGTGTTCTGTGGAAGCC 881
      |||
Db      856 GAGATCGTAGTGCACAGCTTCACCTGCGGCGGAGTTCTTCTACTGCAACAGCACCAG 915
      |||
QY      882 GACCCACCTGCGCCAGCTGGAACAAGTCATCAACAACCTGGAGCTGGCAATCTTC 941
      |||
Db      916 CTGTTCAACGACCTTGGAACAACACCATCGGCCCAACAACCAACGACCATCACCC 975
      |||
QY      942 AAGAGCGCGGCATCAACCTGGCCGAGCTGCCGAGCGGAGTGGGCAATCCATGGGC 1001
      |||
Db      976 CTGCCCCGCGCATCAAGCAAGATCATCAACCGCTGCAGAGAGTGGCAAGGTCATGTAC 1035
      |||
QY      1002 GTGGGCTCGGGGCGCGGCG---GTGCTGTTTCGCAACAACCGCGGTGTATGAGGCGCG 1058
      |||
Db      1036 GCCCCCCCCCATCGCGGCGCAGATCCGCTGCAGCAGCAACATCAACGGCTGTGTGACC 1095
      |||
QY      1059 CTGCGCAAGGCTATGAGCTGTTCAAGGCAAGCGCGCTGCGCGCTGAGCTGAGCGAG 1118
      |||
Db      1096 CGGACGCGCGCAAGGAGATCAGCAACACCAACGAGATCTTCGCGCCCGCGCGCGGAC 1155
      |||
QY      1119 GTGCGCGGCATGAGACGCGCATCAAGAGACCAACATCATGATGAT---GCCGCGCGCGG 1175
      |||
Db      1156 ATGGCGAACAATGGCGCAGCGAGCTGTACAGTAACAAGTGTGAAGATCGAGCCCTG 1215
      |||
QY      1176 TCCAAAGTTTGAAGAGCTGTGAAGCACCGCGCGCGCGCGCGAGGCGCGCGCAC 1235
      |||
Db      1216 GCGGTGGCCCCCAACCAAGGCGCAAGCGCGCGGTGTGCAAGCGCGAGAGCGCGGTGACC 1275
      |||
QY      1236 GGCACCCCGCGCGCGCTGGCCTGGAGCGCGCGCGCGCTTCACCAAGCAAGAGCGAGG 1295
      |||
Db      1276 CTGGGCGCGCAATGTTCTTGGGCTTCTGTGGCGCGCGCGCAACCATGGGCGCCCGAGC 1335
      |||
QY      1296 GGGGCGATCACTGCGCGCTGGCGCGTGGCCCAACGCGCTGGGCAACGCCAAGAAGCTGATC 1355
      |||
Db      1336 CTGACCTCGAACCGTGCA---GGCCCGCAGCTGCTGAGCGGCATGTCAGCAGCAGAAC 1392
      |||
QY      1356 ACCAAGATGCAAGCGCGCGAGGCCAAGTACGACTTTGTGAGATCATGCGCTG 1408
      |||
Db      1393 AACCTGCTGGCGCGCATGAGGCCCAAGCAGCACTGTGCACTGACCGTGTG 1445
      |||
```

Search completed: June 8, 2003, 10:17:54
Job time : 857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:01:16 ; Search time 139 Seconds
(without alignments)
5292.932 Million cell updates/sec

Title: Af012098

Perfect score: 2399
Sequence: 1 ATCTTACATGACACACAAA.....TGTAAGCAGCAGCTTCTGAG 2399

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents NA:*

1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65.4	2.7	44377	2	US-08-804-227C-7
2	65.4	2.7	44377	2	US-08-804-198-1
3	64.6	2.7	50937	4	US-09-428-517-1
4	63.4	2.6	1182	4	US-09-385-028-19
5	63.4	2.6	11604	4	US-09-385-028-13
6	63.4	2.6	15079	4	US-09-385-028-11
7	62.6	2.6	1217	4	US-09-372-422A-11
8	61.8	2.6	735	3	US-09-003-287-7
9	61.8	2.6	8438	1	US-07-945-283-1
10	61.4	2.6	5392	2	US-08-403-852B-1
11	61.4	2.6	5392	3	US-08-510-646B-1
12	61.4	2.6	5392	4	US-09-231-818-1
13	61	2.5	2713	3	US-08-804-439A-13
14	61	2.5	2713	3	US-08-804-439A-13
15	60.4	2.5	4403765	4	US-09-103-840A-2
16	60.4	2.5	4411529	4	US-09-103-840A-1
17	60.2	2.5	2721	6	5215881-2
18	60	2.5	1388	1	US-08-440-856A-9
19	59.4	2.5	1605	4	US-09-124-541-3
20	59.4	2.5	3472	6	5244792-2
21	59.4	2.5	6733	4	US-09-124-541-2
22	59.4	2.5	13842	4	US-09-105-537-30
23	59.4	2.5	36778	4	US-09-105-537-5
24	59.4	2.5	38506	3	US-09-320-878-19
25	58.6	2.5	15872	4	US-09-105-537-1
26	58.4	2.4	4403765	4	US-09-103-840A-2
27	57.8	2.4	1206	4	US-09-372-422A-13

28	57.8	2.4	11220	4	US-09-105-537-32	Sequence 32, Appl
29	57.4	2.4	1755	1	US-08-068-395A-2	Sequence 2, Appl
30	57.4	2.4	1755	1	US-08-464-365-2	Sequence 2, Appl
31	57	2.4	999	2	US-08-761-277A-46	Sequence 46, Appl
32	56.8	2.4	2277	1	US-08-676-967-5	Sequence 5, Appl
33	56.8	2.4	2277	1	US-08-676-974-5	Sequence 5, Appl
34	56.8	2.4	2277	2	US-09-098-487-5	Sequence 9, Appl
35	56.6	2.4	1288	1	US-08-440-856A-9	Sequence 5, Appl
36	56.2	2.3	4257	2	US-08-690-473-1	Sequence 1, Appl
37	56.2	2.3	4257	4	US-09-259-821A-1	Sequence 1, Appl
38	56.2	2.3	4257	4	US-08-843-659-1	Sequence 1, Appl
39	56.2	2.3	12801	1	US-08-458-566A-11	Sequence 11, Appl
40	56	2.3	49377	1	US-08-764-235A-1	Sequence 1, Appl
41	55.8	2.3	987	4	US-09-159-106-12	Sequence 12, Appl
42	55.8	2.3	1516	4	US-09-159-106-10	Sequence 10, Appl
43	55.2	2.3	4411529	4	US-09-103-840A-1	Sequence 1, Appl
44	55	2.3	3468	1	US-07-951-715A-2	Sequence 2, Appl
45	55	2.3	3468	2	US-08-459-448A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-804-227C-7
; Sequence 7, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kohnstoss, Stuart A.
; APPLICANT: Rustock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: .350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:

NAME/KEY: CDS
 LOCATION: 31329..36071
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 36155..41830
 US-08-804-227C-7

Query Match 2.74; Score 65.4; DB 2; Length 44377;
 Best Local Similarity 43.54; Pred. No. 0.00012;
 Matches 502; Conservative 0; Mismatches 641; Indels 12; Gaps 4;

QY 475 CGCGGCGCGCCACACCCCGGAGAGCTGGCCGAGGAGCCCTCCGCGCTTGGCTTTAGC 534
 DB 29120 CGAGAGGCGCGGACCTGTGTGAGAGTGGCCCGCGCTCAACCGGTTGTCCCTCGGCG 29179
 QY 535 AGGTGTTTGAACAGCTGTTTGGCGCGACCTGAACATGAGAGGAGGAGAGAGCTTC 594
 DB 29180 ACCGCGTGTCTGGCTGTGGAGCGGCGCTGGGCGCGCTGTGCTGGCCGACACCGG 29239
 QY 595 TGCACCGCTCAACCGAGACCTGAGAGCCCAACCGCATCTCGACGAGCGCTGCCATGT 654
 DB 29240 TGTCTCCCGCGCTCCGAGAGGCTGTCTACCGCCAGCGCGCTCCGCGCTGT 29299
 QY 655 TCAACAGCTGTGCTCCCGCTGGATGCTATGCTGAGAAATCTTACCGGACCTGATCC 714
 DB 29300 TCTCTACGCGCTTCTACAGCTGTGTGTACCTGTGCTCCGCTCAAGCGCGGAGCGGCTG 29359
 QY 715 CCTACGTGAGACCTGAGAGAGCCCGAGATGATGTCTGGCGCATGTCTTACCTTACC 774
 DB 29360 TGTGTGACCGCGCGCGCGGCGGTGGGATGTGCGCGGTGAGATGCGCGCGCATCTCG 29419
 QY 775 TAGCGGAAAGAGGAGTGGCGCGCAAGAGACT--GGTATGTGTCTCATATGCTCT 831
 DB 29420 GCGCGAGGTGTGTGGCGCGAGCGCGCAAGTGGAGCGCTGTGCGCGCATGAGCA 29479
 QY 832 GCAAGCGCAAGAGCTGAGAGGCTGACCGAGCTGTGTGTGTGAGCGCGACCCGACCC 891
 DB 29480 TCAACGAGACACACTCTGCTCTCTCCGCAACCTTCGACCTTCCGACCGGCTTACCGGAG 29539
 QY 892 TGGCGCAGCTGAGACACGTCATACACCGTGA--GCTGGCGAACATTTTCAAGAGC 948
 DB 29540 CGAGCGGACGTCGCGCGCGAGAGTCTCTGAACTGCTACCAAGAGATGTGTGAGCG 29599
 QY 949 GCGGCACTAACCTGTGCGAGAGTCCCGAGGCGAGTGGGCAATCCAAATGGGGTGGCT 1008
 DB 29600 CTTCTCTGGGCTGTCTCTGCGCGCGCGCTTCTGTGAGCTGGGCAACGAGCTCC 29659
 QY 1009 CGGCGCGCGGCTGTCTGTGCGACCAACCGGCGTGTCAAGAGCGCGCTGCGACG 1068
 DB 29660 GGAACCGCGAGCGGATGCGCGCGAACAACCGGCGTGTCTACCGGCTTTCGACTCA 29719
 QY 1069 CTTATGAGCTTTCACGCGGACCGCGCTGCGCGCTGAGCGTGAAGGAGTCCGCGCA 1128
 DB 29720 ACAGAGCGGAGCGGAGCGACTCGCGCGCTGTGCGGAACTGATGAGACTTTCGCGC 29779
 QY 1129 TGAACGCGATCAAGAGAGCAACATCAATGATGCTGCGCGCGCGCTCCAAATTTGAG 1188
 DB 29780 CCGCGCTGTGCAACCGCTGTCTCTCTCAACCAAGAGTGGCGCGCGCGCG--GACG 29836
 QY 1189 AGCTGTGAAGACACCG 1248
 DB 29837 CCTGTGCGACCATATGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 29896
 QY 1249 CGCTGAGCTTGGAGCGGCGCGCGCTTTCACAGCGAGACGCGAGGCGCGCGCATCAC 1308
 DB 29897 CTTGCGACACCGTACGGGACGGTCTGTGTCAACCGGTGCGACCGGCGCGCTCGGACGCGCA 29956
 QY 1309 TGGCGCTGTGCG 1368
 DB 29957 TCG 30016
 QY 1369 CGGCGAGGCGCAAGTACGACTTTTGTGAGATCATAGCTGCGCGCGC--GGGCTGTGTGG 1425

DB 30017 GCCCGAGCGGAGGCGCGCGCGAGCTGTGCGCGACCTGCGCGCGCTTGGCGCGCTCG 30076
 QY 1426 GCGCGCGCGCGCGCGCGCGCTTCAACGAGCCATACGAGAGAGCGCGCGCGCGCG 1485
 DB 30077 CCACGCTGTGTGCGCTGCGACGCTTCCGAGCGGAGCGCGCTTCCGCGACTTCCGCGCA 30136
 QY 1486 TGTACACCTTGAGAGAGTTCACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1545
 DB 30137 TACCGCGCGATACCGCGCTGAGCGCGCGCTGTGTGCAAGACACCGCGCTTCCGAGCGGCG 30196
 QY 1546 AGCTGTACGACAGTACTTGGAGAGCGCGCTGGGCGCACAGGCGCGCAAGCTGTGCA 1605
 DB 30197 TGTCTCCCGCGCTCACCGCGCGAGATGCGCGCGCTGTGCGCGCGCAAGTGTGAGCGCG 30256
 QY 1606 CCCACTACTGTGCGCG 1620
 DB 30257 CCGTCCACTGTGAGCG 30271

RESULT 2
 US-08-804-198-1
 ; Sequence 1, Application US/08804198
 ; Patent No. 5945320
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgett, Stanley G.
 ; APPLICANT: Kuhstoss, Stuart A.
 ; APPLICANT: Rao, Nagaraja R.
 ; APPLICANT: Richardson, Mark A.
 ; APPLICANT: Rostek, Paul R., Jr.
 ; TITLE OF INVENTION: PLATENOILIDE SYNTHASE GENE
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PAUL R. CANTRELL, 1138
 ; STREET: LILLY CORPORATE CENTER
 ; CITY: INDIANAPOLIS
 ; STATE: IN
 ; COUNTRY: USA
 ; ZIP: 46285
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: Macintosh 7.0
 ; SOFTWARE: Microsoft Word 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/804,198
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CANTRELL, PAUL R.
 ; REGISTRATION NUMBER: 36,470
 ; REFERENCE/DOCKET NUMBER: P9113
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 317-276-3885
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 44377 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 350..14002
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 14046..20036
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 20110..31284
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 31329..36071
 ; FEATURE:

```

NAME/KEY: CDS
LOCATION: 36155..41830
US-08-804-198-1

Query Match
Best Local Similarity 2.7%; Score 65.4; DB 2; Length 44377;
Matches 502; Conservative 0; Mismatches 641; Indels 12; Gaps 4;

QY 475 CCCCCGCGCCACACACCCCAAGCAGCTGAGCGAGGCGCTCCGCGCTGCTTTGAGC 534
DB 29120 CCGAGGCGCCCGGACACCTGCTGAGAGTGGCCCCCGGCTCACCGGTGCGCGGCG 29179
QY 535 AGGTGTTGACACGCTGTTTGGCGCCGACCTGACATCATGAGAGGCGAGGACCTCC 594
DB 29180 ACCGCGTGTCCGCGCTGTGGGAGCGCGGCGCTGGGCGCGCTGTGCGCGACACCGCG 29239
QY 595 TCGACCGCTCAACCGAGGACCTGAGAGCCACCGGACCTCCGACGAGCGCGCTGCCATGT 654
DB 29240 TCGTCCGCCCCCGTCCGAGACGCGCTGTCTTACCGCCAGGCGCGCTGCTCCGCGGTG 29299
QY 655 TCACACGCTGCTGCCCCGCTGAGTGTATGCTGAGAAATCTTACCGGACCTGATCC 714
DB 29300 TCTCTAGCCCCCTACTACGCTGTGTACCTGCGCGGCTCAGGCGCGGAGGAGCGGTGC 29359
QY 715 CCTACGTCGACAGCTGCAAGAGCCCGCAAGATGATGTCGCGGCGCATGTCAAGTCTTACC 774
DB 29360 TCGTGCACCGCGCGCGCGGCGGCGTGCAGATGCGCGGTCAGATGCGCGCGCACCTCG 29419
QY 775 TAGCGGAAAAAGAGGCGATGCGCCCAAGAGACT---GGTATGTGTCTCATATATCCCT 831
DB 29420 GCGCGGAGGTGCTGGGACACCGGAGCGCCCGCAAGTGGAGCGCCCTGCGCGCATGGGCA 29479
QY 832 GCACGCGCAAGCAGTGTGAGAGGCTGACGCGACCTGTTCTGTGTGAGCGCGGACCCGACC 891
DB 29480 TACCGACGACACACTCTGCTCTCTCCGACCTTCGACCTTGGACCGCTTACCGGAG 29539
QY 892 TCGCGCAGCTGACGACGATCATACCAACCGTGA---GCTGGGCAACATTTTGAAGAGC 948
DB 29540 CCGACGCGACGTCCTCCGCGGAGAGTGTCTGTGAATCGCTCACCAAGAGTGTGGAGC 29599
QY 949 GCGGATCAACTGTGCGGAGCTGCCGAGGCGAGTGGGACATTCMAATGGGCTGGGCT 1008
DB 29600 CTTCTCTGCGGCTGCTCCGTCGCGGCGCGCGGCTTCTGTGAGCTGGGCAACGACCTCC 29659
QY 1009 CCGGCGCGGCGGCTGTTTGGGACCAACCGGCGGATCATGAGGCGCGCTGCGACGG 1068
DB 29660 GGGACCGCGAGGAGATGCGCGGAAACACCCCGGAGTGGCTACCGGCGTTGACCTCA 29719
QY 1069 CCTATAGCTGTTCACGAGGACCGCGCTGCGGCTGAGCCTGAGCGAGTCCGCGCA 1128
DB 29720 ACGAGCGCGGAGCCGAGGACCTGCGGCGGCTGCTGGGAGACTGTATGGACCTGTGGCG 29779
QY 1129 TGGACGCAATCAAGAGACCAATCATCATGTGCTCCGCGCTCCAGTTTGAAG 1188
DB 29780 CCGGCTGTGTCACCGCTGCTGCTGTCTGTACCAACGATGCTGCGCGCGCGC---GACG 29836
QY 1189 AGCTGTGAGACCGCGCGCGCGCGCGCGCGGCGCGGACCGCGGCAACCGACCCCGCGCG 1248
DB 29837 CCTTGGGACATCAAGCGAGGCGCGGACACCGGAAAGCTGTGCTTGTACATGCGCGCG 29896
QY 1249 CCGTGTGCTTGGAGCGCGGCGCGGCTTTCACGAGGAGACCGGAGGCGGATCATCAC 1308
DB 29897 CTTGGGACCGCTACCGGAGCGGTCTGTGTACCGGTGGCACTGGGCGCTTGGGACGCGCA 29956
QY 1309 TCGCGCTGGCCTGTGGGCAACGGGCTGGGCAACCGCAAGAGCTGATCATCAAGATGAGG 1368
DB 29957 TCGCCCGGCACTGTGGGAGCGGACCGGCTGCGCGGCTGTGATGCGCGCGCGCGG 30016
QY 1369 CCGGAGGCGCAGTGAAGCTTGTGGAGATCAATGCTGCGCGC---GGGCTGTGTGG 1425
DB 30017 GCGCGGAGCGGAGGCGCGCGGAGGTGTGTGCGGACCTGCGCGCTTGTGGGCGGCTGG 30076
QY 1426 GCGGCGGCGGCGAGCCCGCTTCCACGACCAAGGCAATCAGCAAGGCGGCGGCGCG 1485

```

```

DB 30077 CCACGCTGTCCCTGCGACGCTCTCCGACGCGAGCGCCCTCCGCGACTGTCTGCGGCA 30136
QY 1486 TGTACACCTTGAGCAGAGAGTCCACGCTGCGCGCGGCGGACGAGAACCCGTCATCGCG 1545
DB 30137 TACCGGCGATACCCGCTGAGCGGCTGTGTGTACAGCACCGCGCTCTTCAGACGCGG 30196
QY 1546 AGCTGTACGACAGTACTTGTGAGAGCGCGCTGGGCGACAGGCGGACGAGCTGTGACAGA 1605
DB 30197 TGTGCGCGGCGTCAACCCCGAGCGGAGTGGCGGCTGTGCGGCGCAAGTGTGAGAGCGG 30256
QY 1606 CCGACTACTGTGCGCG 1620
DB 30257 CCGTCCACTGTGAGCG 30271

RESULT 3
US-09-428-517-1/c
; Sequence 1, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Bellach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; EARLIER FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 50937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant DNA
US-09-428-517-1

Query Match
Best Local Similarity 2.7%; Score 64.6; DB 4; Length 50937;
Matches 376; Conservative 0; Mismatches 419; Indels 30; Gaps 3;

QY 857 CCGGACTGTGTTCTGTGTGAGCGCGGACCCGACCTGCGCGAGCTGAGACGTCATCAC 916
DB 3774 CCGCCCTCGGAGGAGTGTTCCTCCGCGGAGGCGGCTGTCTTACGACGCGACAT 3715
QY 917 CACCTGTGAGCTGTGGAAATCTTCAAGAGCGCGGATCATCACTTCCGCGGAGCTGCCGA 976
DB 3714 CACCTGTGCGGTACCGGAGTACGATCGCATCGCGCGCGCGCGGCGGCAACGCTTA 3555
QY 977 GGGCGAGTGGACCAATTCATATGAGGCGCTGCGGCGCGGCGGCTGTGTCGACAC 1036
DB 3654 CGTGAACTCGACGCGCAGAGCCGCTTGGCGCGCTTGGCTTACGCGACGCTGAG 3595
QY 1037 CCGCGGTGTATGAGGCGCGGCTGCGGACCGGCTATGAGTGTTCAGGCGGCGCGCT 1096
DB 3594 CAGCGCGCCATATGCGGCTGCGGCGCGGAGCGACAGAGCTGCGCATCACCGCGCGCT 3535
QY 1097 GCGCGCGCTGAGCGCTTACGAGGTGGCGGCGGATGACGCGCATCAAGAGACCAATCAC 1156
DB 3534 GGGGAAGTGGCCCATATACCTTGAGCTGAGCTGAGTGTCTCCGACGCGCGCGAGAGCG 3475
QY 1157 CATGTGCGCGCGCGCGGCTCAAGTTTGAAGAGCTGTGAAGCAACCGCGCGCGCGG 1216
DB 3474 GATGACCGGG---CGGTGTGATGCGCGGAGCGGAGAGACCGCGGTGTCTTCCGCTA 3418
QY 1217 CCGGAGGCGCGCGGCGGACCGCGCGGCGCTGTGAGCTGTGAGCGGCGCGCGGCTT 1276

```

```

Db      3417 CGACGACGCTCGGAGAGCGGGACCGCACGTGCTCTCCCTGCGCGCGCCAGAGCA 3358
QY      1277 CACACGAGAGACGGGAGGGCGGCAATCACTGCGCGTGGCGGCA----- 1327
Db      3357 CCGATCAGGCGCGTGGCGCGGCTCAACCCAGGACCGTGTGTCTCAACACGGGCTC 3298
QY      1328 -----CGGCTGGGCAACGCCAAGAGCTGATCAACCAATGACGGCGG 1372
Db      3297 CTCCTCACCATTGCTCTGTGTGAGAAAGACCGCGGGTGTCTACCATGTGTGATCCCGG 3238
QY      1373 CGAGCCAGATGACACTTGG---TGAGATCATGAGCTGCGCGGCGGTGTGTGGCGG 1429
Db      3237 CCGAGCGGGCGCGGAGCGACACCGCGCTGTCTTTCGCGAGCGCGACCGCAGAGCGCG 3178
QY      1430 CGGCGGCGACCGCCGCTCCACCGACAGAGCCATCAAGCAAGCGGCGAGCGCGCTGTA 1489
Db      3177 GCTGACCGACGACTTCCCGCGCGACGAGGGGCGAGCGCGTTCGCGGGAGCGCGAGCG 3118
QY      1490 CAACCTGAGACGAAAGTCCACGCTGCGCGCGCGACCGACAGAAACCGCTCATCCGCGAGCT 1549
Db      3117 CTACCGCGGAGTGAATGACCAACTGATCTACCTCGAGGGCATCTACTCCGCTACCGCTG 3058
QY      1550 GTACGACAGCTACCTCGAGAGCGCGCTGGGCAAAAGCGCAGAGCTGCTGCACACCA 1609
Db      3057 GTACGACAGCAGAGCGGCTCCAGCGCGCTTCTCTCTCGCCACGGGTTGTCTACACTC 2998
QY      1610 CTACGTGCGCGCGCGCTGAGAGAGAAAGACGAAAGTGTAGG 1654
Db      2997 CTTCGACTACCGGAGACTGAAGGTGACGGGAGCGCGGAGCGGCGG 2953

```

```

RESULT 4
US-09-385-028-19
; Sequence 19, Application US/09385028
; Patent No. 6232106
; GENERAL INFORMATION:
; APPLICANT: Susan E. Jensen
; APPLICANT: Kwamena A. Aidoo
; APPLICANT: Ashish S. Parakhkar
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic
; Patent No. 6232106
; TITLE OF INVENTION: Acid Biosynthesis
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC
; STREET: The Jennifer Building, 400 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (ERO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/385,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/790,462
; FILING DATE: 29-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: D. Douglas Price
; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 1418/P57452US2
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 39305350
; TELETYPE: RCA 248593 IDEA UR
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1182 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-385-028-19

Query Match      2 6%: Score 63.4; DB 4; Length 1182;
Best Local Similarity 43.5%: Pred. No. 0.00013;
Matches 475; Conservative 0; Mismatches 586; Indels 30; Gaps 3;

QY      405 AGCGCAAGCAGATGTGCTGAGGAGTGTGCTCGGCGCTTCTGTGTGCTATTGCGGAGACC 464
Db      22 AGCGCGCGGAGATTGTGTGTGACACAGGCGCGGTGTGGCTTGGCGAGACGGCGCGAC 81
QY      465 CTGGGCTTGGCGCGCGCGCGCACACCCCAAGACAGCTGGCCGAGAGGCTTCCGCGCTC 524
Db      82 GACTTACCGCTCTGCTGCTTCAACCGCGCGCACCGGTGAGCGCGTCTTCAACCGCTCC 141
QY      525 GCGTTTGAAGAGTGTGTAACAAGCTGTGTGGCGCGACCTGACCATATGAGAGAGGGC 584
Db      142 CGCTTTCGCGGCGCGAGCGTGTGTGTGCGGAGCGGTGTGCGACCGGCGAGCGCGC 201
QY      585 AGCGAGCTGTGCTACCGCTTCAACGAGCACTGAGGCCCAACCGCACTTCGACGAGCGG 644
Db      202 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 261
QY      645 CTGCGCATTTTCAACAGCTGTGCTGCGCGCGCTGATGTCTATGTGTGAATACTTACCGG 704
Db      262 GCGCGGAGAGTGTGCGAGCGCGCTGCGCGCGCGCTGCGCGCGCGCGAGATGTCTG 321
QY      705 GACCTGATCCCTAGCTGAGCGAGCTGCAAGAGCGCGCGCGCATGATGCTGCGCGCATGTG 764
Db      322 ATGCTGCTCAACGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 381
QY      765 AAGTC-----CTACCTACCGGAAAGAGGCGATGCGCCCAAGAGCATGTGTATGTG 818
Db      382 AAGAGCTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 441
QY      819 TCCATATGCTGTGACCGGCAAGAGCTGTGAGGCTGTGACCTGTGTGTGTGTGTGTGTGT 878
Db      442 ACAGCAGACCGCGCGCGCGCAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 501
QY      879 GCGGACCCCAACCTGCGCGAGCTGACGAGCATCATCACCGCTGAGTGTGCGGCAATC 938
Db      502 ATGCGCAAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 558
QY      939 TTCAAGAGCGCGGCAATCACTGTGCGAGCTGTGCGAGGCGAGTGTGGAACAATCCAATG 998
Db      559 GCCAGGAGCGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 618
QY      999 GCGGTGGCTGTGCGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1058
Db      619 CGGACCTTCAACCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 678
QY      1059 CTGCGCAGC-----GCTATGAGCTGTTCAGGCGACGCGCGCTG 1097
Db      679 TTGCGCAAGCGCTGTGCGCGCGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 738
QY      1098 CCGCGCTGTGAGCTGTGAGGAGTGTGCGCGCATGTGAGCGGCATCAAGAGACCAATCATACC 1157
Db      739 GCGCGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 798
QY      1158 ATGTGTGCGCGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1217
Db      799 GAGGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 858
QY      1218 GCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1277
Db      859 GTTAACTCTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 918
QY      1278 ACCAGCGAGAGCGGAGGCGCGCATCACTGTGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 1337
Db      919 GCCATGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 978

```

QY 1338 AACGCCAAGAGTATCAACCAAGATGACGGCCGGCCGAAGTACGACTTTGTGAG 1397
DB 979 CCCTTCGGAGAGTCAAGGTCATCCGCCAAGAGCCCGGGGCGACAGGCCGACGCG 1038
QY 1398 ATCATGAGCTTGTCCCGGGGCTGTGGGCGGGGCGGCGCCAGCCCGCTCCACCGACAG 1457
DB 1039 CTGGGGGCGCGCTCCCGGAGCATCTGGGGGCGAGAGTGATCATCGGAGTCACTC 1098
QY 1458 GCCATCAGCA 1468
DB 1099 GCCATCGCGA 1109

RESULT 5
US-09-385-028-13
; Sequence 13, Application US/09385028
; Patent No. 6232106
; GENERAL INFORMATION:
; APPLICANT: Susan E. Jensen
; APPLICANT: Kwamena A Aidoo
; APPLICANT: Ashish S. Paraskar
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic
; Patent No. 6232106
; TITLE OF INVENTION: Acid Biosynthesis
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC
; STREET: The Jennifer Building, 400 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/385,028
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/790,462
; FILING DATE: 29-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: D. Douglas Price
; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 1418/P57452US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 39305350
; TELEX: RCA 248593 IDEA UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11604 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-385-028-13

Query Match 2.6%; Score 63.4; DB 4; Length 11604;
Best Local Similarity 43.5%; Pred. No. 0.00022;
Matches 475; Conservative 0; Mismatches 586; Indels 30; Gaps 3;

QY 405 ACGCGCAACACCTCTGCTGAGTGCAGTGCCTCCGCGCTTCTGTGCTATTGCGAGACC 464
DB 5884 ACGCCCGGGGATTCTGTGTGACACACGGCGCCGTGGGCTGCGCCAGCGCGCGAGC 5943
QY 465 CTGGGCTGGCGCGCGGCGCCACACCGCCCAAGCAGTGCAGAGGCGCTCCCGCGCTC 524
DB 5944 GACTTACCGTCTCTGCTTCACCGCGCGCGCCAGCCAGCTGAGCGCGCTTTCACCGCTCC 6003

QY 525 GGCCTTGACAGAGTGTTTGACAGCTGTGTCGCGCCGACCTGACCATCATGAGAGAGCGC 584
DB 6004 CCCTTGCGGGGCGGAGCGCTCGGTGTGTCCGGAGAGCGGTGGCCGACGGGAGGGCGCC 6063
QY 585 AGCGAGCTGCTGACCGCTTACCGAGACCTGAGAGGCCACCGGACCTCCGACGAGCGC 644
DB 6064 GGTGTGTGTGTCTGTGCTCCGCAACCGGAATGTCCGACCGGCTGAGAGGCGAGAAC 6123
QY 645 CTGGCCATGTTACACAGCTGTGCGCCCGCTGGATGCTATGCTGAGAAATCTTACCGC 704
DB 6124 GCGCGGAGGTGCGAGGCGCTGCGCGCCCTCGGCTCGAGGCTCGAGGCGAGATGCTG 6183
QY 705 GACTGATCCCTTACGTGAGCAGCTGCAAGAGCCCGCCAGATGATGCTGCGGCGCATGCTC 764
DB 6184 ATGCGCTCACCGGGGTGATTCGGCGGCGAGTACCGATGAGAGCATCCGGAGACCTC 6243
QY 765 AAGTC-----CTACTAGCGGAAAGAGGCGATGCGCCAAAGACATGTCATGTCG 818
DB 6244 AAGACGCTGAGTGGCGCGCGGGGAGGGCGGCTTGACCGCGCGCGCCGACATCATG 6303
QY 819 TCATCATGCTCTGACCGGCAAGCATGTGGAGCTGACCGGAGCTGTCTGTGTGAC 878
DB 6304 ACGACCGACACCGCGCCCAAGAGGTCCGGGTCAAGCTGCGGGCGGACCTCTGTGGC 6363
QY 879 GCGAACCACCTGCGCAGCTGAGCACAGTCACTACACCGTGGAGCTGGGCAATC 938
DB 6364 ATGCGCAAGGCGTCCGACATGCTGAGCCCGACATGCGAC---GCTGCTGACTTCTTC 6420
QY 939 TTCAGAGACCGCGCATCACTGTGCGAGCTGCCGAGGGGAGTGGAGCAATCCAAATG 998
DB 6421 GCGACGAGCGCGCGCTGAGCCCGCGAGCGACCGCTCTTCGCGGGGTATGAC 6480
QY 999 GCGTGGGCTGGGCGCGCGCTGTGTGGACACCGCGGTCTATGAGAGGGCGC 1058
DB 6481 CGACCTTCAAGCGGTCAAGATGACACCGACACTTCAACGACGACACGGCGGCTGTC 6540
QY 1059 CTGGCGACG-----GCTATGAGCTTCTTACGGGGACGCGCTG 1097
DB 6541 TTGCGCAAGGCGCTGCGGGCGAGTGCAGCGCCGGGAGTTGAGAGAGGCGCTGCACAGC 6600
QY 1098 CCGCGCTGAGCTGACGAGTGGCGGCGCATGAGCGCATGAAGAGACCAATCATC 1157
DB 6601 GCGGCGCTGCGCTGTCAAGGACATCGGAGCGAGCGGCGCGGCTCAAGCTGATC 6660
QY 1158 ATGTGCCCCCGCGCGGTCCAAAGTTTGAAGAGCTGTGAACGACCGCGCGCGCGC 1217
DB 6661 GAGGTCAAGTCAACCGCGCGCGCGAGCGACCGCCAGGCCAAGCGGTCTGGCAAGATCCTC 6720
QY 1218 GCGAGGCGCGCGCGACGCGACCGCCCGGCGCTGAGCTGGGAGCGCGCGCGCTTC 1277
DB 6721 GTCACTCCCGCTGTGTGAAGACCGCGGTGACCGCTGACCGCCAACTGGGCGCGGCTC 6780
QY 1278 ACCAGGAGGAGCGGAGGCGGCGCATCACTGCGGTGCGCGCTGCGCAACGCGCTGGC 1337
DB 6781 GCATGCGCATGGGAGAGTGTGCGAGCACACCGACATGACCAAGAGGCGGTGAGATC 6840
QY 1338 AAGCGCAAGAGTGTATCAACAGATGACGCGCGCGGAGCCAGTGAAGCTTTGTGAG 1397
DB 6841 CGCTTGGGAGAGTCAAGGTCTATCGCGCGAAGCGCGGGGAGCACGAGCGCGAGCGC 6900
QY 1398 ATCATGAGCTGCGCGCGGCGTGTGTGGCGGCGGCGCGACCGCGCTCCACCGAGAG 1457
DB 6901 CTGGGGCGCGCTGCGGAGCATCTGCGGGGCGAGAGGTGTATCGGATCGACCTC 6960
QY 1458 GCCATCAGCA 1468
DB 6961 GCCATCGCGA 6971

RESULT 6
US-09-385-028-1
; Sequence 1, Application US/09385028

TYPE: DNA
ORGANISM: Zee mays
FEATURE:
NAME/KEY: CDS
LOCATION: (110)...(974)
US-09-372-422A-11

Query Match
Best Local Similarity: 48.7%; Pred. No. 0.0019;
Matches 170; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

1004 GGGCTGGGCGCGCGCTGCTGTTGGCACCACCGCGGTGTATGAGCGCGCTGCG 1063
133 GGGGCGCGAGCGCGGAGTTCTCGGCAAGAGTACACCGACCTCCCGCGCGCGCT 192
1064 CACGCGCTATGAGCTGTTCAAGGAGACGCGCGCGCGCTGAGCGCTGAGCGAGTGC 1123
193 GATCGACGCGGAGAGTCTACCCAGTGTACTGTACCGCGCGCTATCGCCGATTCAT 252
1124 CGGCGATGAGCGGCGATCAAGGAGACCAACATCACCATGTTCCCGCGCGCGGTCGAATT 1183
253 CGCGACGCTGCTCTTCTCTTACATCACCGTGACACCGTATCGGGTACAGCACGAGAC 312
1184 TAGAGAGCTGTCTAAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1243
313 GAGCGGCTGCGCGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 372
1244 CGGCGCGCTGCGCGCTGCGGAGCGCGCGCGCTTACACGAGGAGAGCGGCGCGCAT 1303
373 TTGGGCTTGGCGCGCATGATCTTTCATCTCTGCTACTGACCGCGCGCATCTCAGGTG 432
1304 CACACTGCGCGTGCCTGCGCGCAACGCGCGCTGCGCGCAACGCGCGAGCTG 1352
433 GACATCAACCGCGCGCTGACCTTTCGCGCTTCTCGGCGCGAGAGGTG 481

RESULT 8

US-09-003-287-7
Sequence 7, Application US/09003287
Patent No. 6096947
GENERAL INFORMATION:
APPLICANT: Jayne, Susan
APPLICANT: Barbour, Eric
APPLICANT: Meyer, Terry
TITLE OF INVENTION: METHODS FOR IMPROVING TRANSFORMATION EFFICIENCY
FILE REFERENCE: MOPAT MOCAM
CURRENT APPLICATION NUMBER: US/09/003,287
CURRENT FILING DATE: 1998-01-06
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 735
TYPE: DNA
ORGANISM: Myrothecium verrucaria
FEATURE:
NAME/KEY: CDS
LOCATION: (01)...(732)
US-09-003-287-7

Query Match
Best Local Similarity: 45.5%; Pred. No. 0.0024;
Matches 305; Conservative 0; Mismatches 357; Indels 9; Gaps 2;

875 GAGCGCGGACCGGCGCGCTGCGGAGCGGACGTCATACGACCGTGGAGCTGGGCA 934
57 GGCATGCTGAGCTCCTCGGCAAGCTCGCGAGCTAGCTCTTACTCCCTGGAGGACAT 116
935 CATCTTCAAGAGCGCGGCGATCACTGCGCGAGCTCCGAGGCGGAGGAGCAATCC 994
117 CGGCTTCCAGCGCGGCGGAGAGCTCTGCGCGAGGCGGCGGAGGCGGCTTCAAGGCGCGCT 176
995 AATGGCGTGGCTCGGCGCGCGGCTGCTTTCGAGCACCGCGCGGTCTCATGAGGC 1054

177 CTCGCCGAGACCTACACCACTCCATGCGCGCTGTTCTAATGAGGACCGGTATGCGCAG 236
1055 GGGCGTGGCGCAC-----GGCTATGAGCTGTTACGAGGAC---GGCGTGGCGGCGCT 1105
237 GAGGCTCTCTCCGAGGACAGGCCAAGGACCTTCTCCGCTGACCTGAGCCCTCAGCTGCT 296
1106 GAGCTGAGCGAGGTGCGCGGATGAGCGGATGAGCGGATCAAGAGAGACCAATCAGCATGTGCG 1165
297 CTTCCAGAGAGTGGGACCGCGGAGGCTTACTTACCTCACCGGATGCTTTCAGCAT 356
1166 CGGCGCGGCTCCAAATTTGAGAGCTGTTGAAGACCGCGCGCGCGCGCGCGCGCGAGGC 1225
357 CTACGCGCGGATCAAGCGAGTGAAGGCTCTCAAGGCTCTGCTCTCCACGACGAGCG 416
1226 CGCGCGGACGAGGACCGCGCGCGCGCGCGCTGAGCGAGGCGGCGCGGCTTCCAGCGCA 1285
417 GAGGCGGCTGGCGGAGGCGGATATCCGCGACGAGAGCTGCGGCTGAGAGGCGCAATCAC 476
1286 GAGCGGCGAGGCGGCGATCACTGCGCGTGGCGCTGAGCGGCTGAGCGCAAGCGCA 1345
477 CTTCCTGCGGCGAGCTATCCAGCTCGCACCGCTCTACGACCAAGTGGCGGCTTACGAGCG 536
1346 GAGCTGATCAACAGATGAGCGCGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCG 1405
537 CATGACGACTTCTGCTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 596
1406 CTGCGCGCGGCGGCTGCTGAGGCGGCGGCGGCGGCGGCGGCTGCGCGGCGGCGGCGG 1465
597 CGCGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 656
1466 GCGAGAGCGGCGGCGGCGGCTGTAACCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1525
657 GCGGTGTCACACACCAACCACTCCGCGAGTTGAGCAAGGAGGAGGAGGAGGAGGAGGAGG 716
1526 CGAGAACCGCT 1536
717 GATGAGCGCT 727

RESULT 9

US-07-945-283-1/c
Sequence 1, Application US/07945283
Patent No. 5352596
GENERAL INFORMATION:
APPLICANT: Cheung, Andrew K.
APPLICANT: Wesley, Ronald D.
TITLE OF INVENTION: Pseudotables Virus Deletion Mutants
TITLE OF INVENTION: Involving The EPO and ILT Genes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis P. Ribando
STREET: 1815 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4011 ext. 513
TELEFAX: 309-685-4128
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

Db 2137 CTGCGCTTCCGCGCAATCGACACGCTATCCACGCCGCAACCCGCGCATGTTCCGCG 2196
Qy 966 GACCTGCCGAGGCGGAGTGGGACATCCAAATGGGCGTGGGCTCGGGCGCGCGTGGT 1025
Db 2197 GGCCTGGGCGCGTTCTTCGAGGGAGATGGGCGACCTGGGCGCGCGGCTCGCAAGGTGGT 2256
Qy 1026 TTGCGACCAACCGCGGTCTCATGAGCGCGCGCTGGCAACGCGCTATGAGCTTTACG 1085
Db 2257 ATGGGCACTGGGCGCGGTGTATCGCGGTGCGGCGGTCTCTTCATGTCACAC 2316
Qy 1086 GCGACCGCGCTCGCGCGCTGAGCGCTGAGGAGTGGCGCGCATGACCGCATCAAGAG 1145
Db 2317 CCTTTGGGCGCGTGGCGGTGTGTGTGTCTCTGCGCGCGCTGGCGCGCGCTTTCTTC 2376
Qy 1146 ACCAATCATCATGTGTCGCGCGCGCGGTCCAAATTTAGAGAGCTGTGAAAGCAACG 1205
Db 2377 GCTTTGCGCTACGTGATGCGGCTGCAAGCAACCCCATGAAGGCGCTGTACCGCTAAC 2436
Qy 1206 GCTGGCGCGCGCG 1265
Db 2437 ACCAAGAGCTCAAGAACCCGCAACCGCGAGCGCGTCCGCGGAGCGCGAGAGGCGCG 2496
Qy 1266 GCGCGCGCTTACACGACGAGAGAGGCGCGCATCACTGCGCGGTGGCGCGCG 1325
Db 2497 GACTTTACGAGCGCAAGCTAGCGCGAGCGCGGAGATGATACGTACATGCGCTGTG 2556
Qy 1326 AACGCGCTGGGCAACGCGCAAGAGCTGATCAACAGATGAGCGCGCGCGCAAGTAC 1385
Db 2557 TCTGCCATGAGCGCGACGAAACAGAGCGCAAGAGGCAAGAGCGCGGTGTACG 2616
Qy 1386 GACTTTGTGAGATCATGCGCTGCGCGCGCGCGCTGTGTGGCGCGCGCGCGCGCG 1445
Db 2617 GCCAAGTCAACGACATGTGATGCGCAAGCGCGCAACCACTACCGCAAGTTCC 2676
Qy 1446 TCCACCGAC 1454
Db 2677 AACAAAGAC 2685

RESULT 14

US-08-720-229-13
Sequence 13, Application US/08720229
Patent No. 602342
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Boesch, Marix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,229
FILING DATE: 26-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO.: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2713 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-720-229-13

Query Match

2.5%; Score 61; DB 3; Length 2713;
Best Local Similarity 43.2%; Pred. No. 0.00049;
Matches 289; Conservative 0; Mismatches 380; Indels 0; Gaps 0;

Qy 786 AAGGCGATCCCGCCAAAGCAATGTCATGTGTCTCATCATGCTTGCACGCGCAAGCAG 845
Db 2017 AACATCACCATCTGAGAGATCACAGATTGTTCCTCCCTGAGGTATACCCGCAAGG 2076
Qy 846 TCGAGGCTGACCGGACATGCTTCTGTGTGAGCGCGCAACCCGACCTGCGCGAGTGCAG 905
Db 2077 ATCAAGACAGCGCGCTGCTGTGATACAGAGGTTCAGCGCGCAACCAAGTGCAGAGC 2136
Qy 906 CAGTCATCACCAACGCTGAGCTGGGCAACATTTCAAGAGCGCGCATCAACCTGCGC 965
Db 2137 CTGCGCTTCCCGACATGCAAGGTCATCCAGCGCGAGCGCAACGCGCGCATGTTCCG 2196
Qy 966 GAGCTGCCGAGGCGAGTGGCAATCCAAATGGGCGTGGCTCGGCGCGCGCGCTGTG 1025
Db 2197 GGCCTGGGCGCTTTCTTCGAGGGAGATGGCGACTGGGCGCGCGGTGCGCAAGGTGG 2256
Qy 1026 TTGCGACCAACGCGGCGGTTCATGAGCGCGCTGGCGCGCGCTATGAGCTGTTCAG 1085
Db 2257 ATGGGCACTGGGCGCGCGCTGTATGCGCGCGGTGTGCGGCTGTCTCTTCAATGTCAAC 2316
Qy 1086 GGCACCGCGCTGCGCGCTGAGCGTGAAGTGGCGGAGTGGCGGAGTGAAGCATCAAGAG 1145
Db 2317 CCTTTGGGCGCGTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2376
Qy 1146 ACCAATCATCATGTGTCGCGCGCGGTCCAAATTTAGAGAGCTGTGAAAGCAACG 1205
Db 2377 GCTTTGCGCTACGTGATGCGGCTGACGAGCAACCCCATGAAGGCTGTACCGCTAAC 2436
Qy 1206 GCTGGCGCGCGCG 1265
Db 2437 ACCAAGAGCTCAAGAACCCGCAACCGCGAGCGCTCGGAGGCGCGAGAGGCGCG 2496
Qy 1266 GCGCGCGCTTCAACGACGAGAGCGCGAGGCGCGCATCACTGCGCGTGGCGCG 1325
Db 2497 GACTTTGAGAGCGCAAGCTAGCGCGAGCGCGGAGATGATACGTATATGCGCTGTG 2556
Qy 1326 AACGCGCTGGGCAACCGCAAGAGCTGATCAACAGATGACAGCGCGCGCAAGGTAC 1385
Db 2557 TCTGCCATGAGCGCGCAAGCAACAGCGCAAGAGGCAAGAGCGCGCTGTCTACG 2616
Qy 1386 GACTTTGTGAGATCATGCGCTGCGCGCGCGCTGTGTGGCGCGCGCGCGCGCGCG 1445
Db 2617 GCCAAGTCAACGACATGTGATGCGCAAGCGCGCAACCACTACCGCAAGTTCC 2676
Qy 1446 TCCACCGAC 1454
Db 2677 AACAAAGAC 2685

RESULT 15

US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6293328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM


```

; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
; US-09-103-840A-2
```

```

Query Match      2.5% Score 60.4; DB 4; Length 4403765;
Best Local Similarity 48.8%; Pred. No. 0.0041;
Matches 163; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 997 TGCGCGTGGGCTCGGCGCGCGCGCTGCTTCGCGACACCGCGGCTGTCTATGAGGCG 1056
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2059204 TGCGCGGTGTGGCGGTGAGGCTGTGCGATTCTTGCGGACCGCCCGCGGCGCGGTG 2059145

QY 1057 CGCTGGCGACGCGCTATGAGCTGTTCAAGGCGACGCGCGCTGCGCGCTGAGCG 1116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2059144 GTCCCGGCGGGCGGCTGTTGCGGTGCGTGGGCGCGGCGCGCGGAAATCG 2059085

QY 1117 AGGTGCGCGCGCATGAGCGCATCAAGAGACCAATCATCATGTTGCCCGCGCGCGGT 1176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2059084 GATTGCTCGGGAACGCGGCGCGCGGCGGCTCGCGGCTCGCTGCGGCGACG 2059025

QY 1177 CCAAGTTTGAGAGCTGTGAAGACCGCGCGCGCGCGCGAGGCGCGCGCGACG 1236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2059024 GCGGTGCGCGCGCGCGCGCGGCTGCGGCTCGGCTCACTACCGCGGTGCGCGCGGCG 2058965

QY 1237 GCAACCCCGCGCGCGCTGAGCTGAGACGCGCGCGCGGCTTCACTACGAGAGACGCGAGG 1296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2058964 GCAACGCGCGCGCTGTGTGTAGGCGCGCGCGCGCGCGCGCGCACTCGCGCGGTG 2058905

QY 1297 GCGGCATCACTGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2058904 GCGCTACCGGGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2058871
```

Search completed: June 8, 2003, 13:31:28
Job time : 2876 sec

This Page Blank (uspto)

QY 702 CCGGACCTGATCCCTTACCTGAGCAGCTGCAAGACCCCAATGATGCTGGCGGCATG 761
DB 736 GCCCAGCTGCAACATCAGCGCGGAGAGAGTGAACAACCTTGAAGAGATCGTGAACAA 795
QY 762 GTCAAGTCTTACCTAGCGGAAAGAGGAGCATCGCCCAAGAGCATGATGATGTC 821
DB 796 CTGCAAGCCCAAGTTCGGCAACAAGACCATGCTTTCAAGCAGAGAGCGCGCGACCC 855
QY 822 ATCATGCTTGCACCGCGCAAGCAGTGGAGGCTGACCGCGCATGCTTCTGTGAGCGC 881
DB 856 GAGATCGTATGACAGCTTCAACTGGCGGAGTCTTCTTACTGCAACAGCACCCAG 915
QY 882 GACCCCACTGGCCAGCTGAGCAAGTATCAACCAAGTGGAGCTGGCAATCTTC 941
DB 916 CTGTTCAACAGCATCTGGAACAACCATATGCCCCCAACAACAGGACATATAC 975
QY 942 AAGAGCGCGCATCAACCTGCGGAGCTGCCCCGAGGCGATGGAGCAATCCATGGC 1001
DB 976 CTGCGCTGGCGCATCAAGCAGATCATGACCGTGGAGGAGTGGCAAGGCGCATGTAC 1035
QY 1002 GTGGGCTGGGCGCGCGC--GTGCTGTTGGCAACAACCGCGGTGTATGAGGCGCG 1058
DB 1036 GCCCGCCCATCGCGCGCAGATCCGCTGAGCAGCAATCACCGCGCTGTGACCC 1095
QY 1059 CTGCGCAGCGCTTATGAGCTGTTACAGGAGACCGCGCTGGCGGCTGAGCGAG 1118
DB 1096 CGGACGCGCGCAAGAGATCAGCAACACCGAGTCTTCCGCCCGCGCGCGAG 1155
QY 1119 GTGCGCGCATGAGACGCGCATCAAGAGACCAACATCACTATGTT--GCCCGCGCGG 1175
DB 1156 ATGCGGACAACTGGCGCAGAGCTGTATACAGTACAGGAGTGTGAAGATCGAGCCCTG 1215
QY 1176 TCCAACTTTGAGAGAGCTGTGAAGACACCGCGCGCGCGCGCGCGCGCGCGCAC 1235
DB 1216 GCGGTGCGCCCCCAAGGCGCAAGCGCGCGTGTGAGCGCGCAAGCGCGCGTGAAC 1275
QY 1236 GGCACCGCGCGCGCGCTGGCGCTGGAGACGCGCGCGCGGCTTCAACAGCAGAGCGGAG 1295
DB 1276 CTGGGCGCCATGTTCTGGGCTTCTGGGCGCGCGCGCGCGCACCATGGCGCGCGAGC 1335
QY 1296 GCGCGCATCACTGCGCGCTGGCGCTGGCGCAAGCGGCTGGCAACCGCAAGACTGATC 1355
DB 1336 CTGACCTGACCGTGA---GGCGCGCAGCTGTGAGCGGATCGTGAAGAGCAAG 1392
QY 1356 ACCAAGATGACGCGCGCGGAGGCGCAAGTACACTTTGTGAGATCATGCGCTG 1408
DB 1393 AACCTGCTGCGCGCATCGAGCGCAAGCACCTGCTGAGCTGACCGTGTG 1445

RESULT 2
US-09-476-242-4
; Sequence 4, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: MARTIN, Eric
; APPLICANT: HARTOG, Karin
; FILE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2316
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Val120-11e201
US-09-476-242-4

Query Match 2.9%; Score 69.8; DB 10; Length 2316;
Best Local Similarity 44.3%; Pred. No. 4e-09;

Matches 422; Conservative 0; Mismatches 522; Indels 9; Gaps 3;
QY 462 ACCCTGGGCTTGGCG 521
DB 502 ACCCAAGGAGTCCG 561
QY 522 CTGGGCTTGGAGAGGTTTACAGCGCTGTTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 581
DB 562 GAGGCGTGTGATCCGAGCGAGAACTTACCGACACAGCCCAACAGCCGCAAGAGCATC 621
QY 582 GCGAGGAGCTGCTGACCGCGCTGACCGAGCATCTGAGGCGCGCGCGCGCGCGCGCGCG 641
DB 622 CTGAAGAGAGGTGAAGATCACTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 681
QY 642 CCGCTGCGCATGTTACAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 701
DB 682 ACCATCG 741
QY 702 CCGGACCTGATCCCTTACCTGAGCAGCTGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCG 761
DB 742 GCCCATGCAACATCAGCGCGCGAGAGTGAACAACCTTGAAGAGATCTGTACCAAG 801
QY 762 GTCAAGTCTTACCTAGCGGAAAGAGGAGCATGCGCGCAAGAGCATGATGTC 821
DB 802 CTGCAAGGCGCGAGTGGCGCAACAAGACCATGTTCAAGCAGAGAGCGCGCGCGCGCGCG 861
QY 822 ATCATGCTTGCACCGCGCAAGCAGTGGAGCTGAGCGCGCATGGTCTGTGTGAGCGCG 881
DB 862 GAGATGATGATCAGACGCTTCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 921
QY 882 GACCCGACCTGGCGAGCTGGAGCAAGTATCACCAGCGTGGAGCTGGCGCAACATCTTC 941
DB 922 CTGTTCAACAGCATCTGGAACAACCATATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 981
QY 942 AAGAGCGCGGAGATCAACTGCGCGAGCTGCCGAGCGGAGTGGAGCAATCCATGGCG 1001
DB 982 CTGCGCTGCGCATCAAGAGATCATCAACCGCTGGCAGAGAGTGGGCGCAAGGCGCATGTAC 1041
QY 1002 GTGGGCTGGGCGCGCGC--GTGCTGTTGGCAACAACGCGCGGTGTATGAGAGCGGCG 1058
DB 1042 GCCCGCCCATCGCGCGCGAGATCGGCTGACAGAGCAATCAACCGCGCTGTGACCC 1101
QY 1059 CTGCGCAGGCTTATGAGCTGTTCAAGGCGCGCGCTGCCGCTGAGCGCTGAGCGAG 1118
DB 1102 CCGGAGCGCGCAAGAGATCAGCAACACCATCGAGATCTTCCGCGCGCGCGCGCGCG 1161
QY 1119 GTGCGCGGATGAGCGCATCAAGAGCAACATCAGCATGTT--GCCCGCGCGCGG 1175
DB 1162 ATGCGGACAACTGGCGCAGCGAGCTGTACAAATGTAAGAGTGAAGATCGAGCCCTG 1221
QY 1176 TCCAACTTTGAGAGAGCTGTGAAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1235
DB 1222 GCGGTGCGCCCCCAAGGCGCAAGCGCGCGTGTGCAACCGCGCGAGAGCGCGCGTGAAC 1281
QY 1236 GGCACCGCGCGCGCTGGCGCTGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1295
DB 1282 CTGGGCGCATGTTCTGGGCTTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 1341
QY 1296 GCGGCGATCACTGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1355
DB 1342 CTGACCTGACCGTGA---GGCGCGCAGCTGAGCGGCGCGCTGAGCGCGAGAGC 1398
QY 1356 ACCAAGATGACGCGCGCGCGAGGCGCAAGTACACTTTGTGAGATCATGCGCTG 1408
DB 1399 AACCTGCTGCGCGCATCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1451

RESULT 3
US-09-476-242-8
; Sequence 8, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan

Oy	1002	GTGGGCTTGGGCGCCGCGC---GTGCTGTTGGGACACACCGGCGGTGATATGAGGCGCG	1058
Db	1042	GCCCCCCCACATCCGGGCGCAAGATCCGCTGCAAGCAACATACCGGCTCTGTGTAC	1101
Oy	1059	CTGGCGACAGCGCTTATGAGCTGTTCACGGGACGCGCTGCCGCTTGAAGCTTAGCGAG	1118
Db	1102	CGCGACGGCGGCGAAGGATATCAGCAACACACCGAGATCTTCGCGCCCGGCGGGCGAC	1161
Oy	1119	GTCGGCGGCATGACCGCATCAGGAGACCAACATCACATGT---GCCCGCCCGGG	1175
Db	1162	ATGCGCGCAACTGGCGCGAGCGAGCTGTACAAAGTCAAGGGTGGTAAATGAGCCCTG	1221
Oy	1176	TCCAAATTGAGAGAGCTGCTGAAAGCACCGCGCGCGCGCGCGCGAGGCGCGCGAC	1235
Db	1222	GCGGTGGCCCCCACCAGGCCAAGGCGCGCGCTGTGTGACGCGGAGAACCGCCCTGAC	1281
Oy	1236	GGCACCCCCCGGCGCTGTGGACGCGCGCGCGGCTTCACAGGAGGACGGCAGG	1295
Db	1282	CTGGGCGGCATTTCTTGGGCTTCTTGGGCGCGCGCGGACCATGAGGCGCCCGCAGC	1341
Oy	1296	GCGGCATCACCTGGCGGCTGGCGCTGGCCAAAGGCTGGGGAAAGCCAAAGTGTATC	1355
Db	1342	CTGACCCCTGACCGTGCAC---GGCCCGCGAGCTGCTGAGCGGGATGTGCAGGACGACAA	1398
Oy	1356	ACCAAGATGCGAGCGCGGCGAGGCCAAAGTACGATTTGTGAGATCATGAGCCTG	1408
Db	1399	AACCTGCTGCGGCATCGAGGCCCAAGCAGCACTCTGCTGAGCTACCTGTG	1451

RESULT 5
US-09-476-242-6
; Sequence 6, Application US/09476242
Date of No. US000001466201

```

1  GENERAL INFORMATION:
2  APPLICANT: BARNETT, Susan
3  APPLICANT: HARTOG, Karin
4  APPLICANT: MARTIN, Eric
5  TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
6  FILE REFERENCE: 1605.002
7  CURRENT APPLICATION NUMBER: US/09/476,242
8  CURRENT FILING DATE: 1999-12-30
9  NUMBER OF SEQ ID NOS: 26
10 SOFTWARE: PatentIn Ver. 2.0
11
12  SEQ ID NO 6
13  LENGTH: 2328
14  TYPE: DNA
15  ORGANISM: Artificial Sequence
16  FEATURE:
17  OTHER INFORMATION: Description of Artificial Sequence: Lys121-Val200
18
19  US-09-476-242-6

```

Query Match	2.9%	Pred. 69.8%	DB 10%	Length 2338
Best Local Similarity	44.3%	Pred. No. 48-09		
Matches 422; Conservative	0;	Mismatches 522;	Indels 9;	Gaps 3

[illegible]

Db	748	GCCCACTGCGAACATCAGCGCGCGAGAAAGTGGAAACAACACCTTGAAGCAGATCGTGACCAG	807
Qy	762	GTCAAGTCTTAACCTAGCGGAAAAAGAGGCGCATGCGCCAAAGGACATGTGTCATGTGTCC	821
Db	808	CTGCAGGCCCAAGTTCCGCAACAAGACCATGTTCTTCAMGACGAGCAGCGCGGACCC	867
Qy	822	ATCATGCCCTTGACACGCGCGCAAGCAGTGGAGAGCTGACCGGACACTGGTTCTGTGTGAGCGC	881
Db	868	GAGATCGTGTATGCACACACTTTCACACTGGCGCGCGGAGTTCTTACTCTCAACAGCACCCAG	927
Qy	882	GACCCCAACCCCTGCGCGCAGCTGAGCAACGTCATCCACCGTGAAGCTGGGCAACATCTTC	941
Db	928	CTGTTCAACAGCACCTTGGAAACAACACCATGGGCCCAACAACACCAAGCACCATCAC	987
Qy	942	AAGAGGCGCGGCATCAACTTGCGCCGACCTGCGAGGGCGAGTGGGAAATCCAAATGGGC	1001
Db	988	CTGCCCTTGCCCATCAAGCAGATCATCAACGCGCTGGAGAGAGTGTTGGCAAGGCATGTAC	1047
Qy	1002	GTGGGCTTCGGGGCGCGCG---GTGCTGTGTGGGACCAACCGGCGGTGTATGAGAGCGGCG	1058
Db	1048	GCCCCCCCCCATCCGCGGCGAGATCGCTGACGACGCAACATCACCGGCTGTCTGTAC	1107
Qy	1059	CTGCGCAACGCGCTTATGAGCTGTTTACACGGGCAACGCGCTGCGCGCTTGAGCTTAGCGAG	1118
Db	1108	CGCGACGGCGGCAAGAGATATAGCAACAACACCGAATCTTCGCGCCGCGCGCGCGAC	1167
Qy	1119	GTCGCGGCGCATGACCGGCATCAAGAGAACCAATCACATGTT---GCCCGCGCCGGG	1175
Db	1168	ATGCGCGCAACTGSGCGGACGACCTGTATCAATATCAAGAGTGTAAGATATGAACCCCTG	1227
Qy	1176	TCCAAGTTTGAAGAGCTGCTGAACACACGCGCGCGCGCGCGCGCGAGAGCGCGCGCAC	1235
Db	1228	GCGGTGGCCCCCACCAAGGCCAAGCGCGCTGTGTGACGCGCGAGAAAGCGGCGCTGAC	1287
Qy	1236	GGCACCCCCCGGAGCGCTGTGGACGCGCGCGCGCGCTTACACAGCGAGACGCGAG	1295
Db	1288	CTGGGCGGCATGTTCTTGCGGCTTCTGTGGCGCGCGCGCACACCATGGGCGCGGAC	1347
Qy	1296	GCGGCGATCACCTGCGCGTGCGCGTGCGCCCAACGGGCTGGGCAACGCCAAGAACTGATC	1355
Db	1348	CTGACCTTGACCTGTGCA---GGCGCGCAGCTGTGAACGGCACCTGTGTACACACAGAAC	1404
Qy	1356	ACCAAGATGACGCGCGCGGAGGCGCAAGTATGACATTTGTGAGATCATGGCTTG	1408
Db	1405	AACCTGTGCGGCGCATATGAGGCCACGACGACCTGTCTGACGCTGACCGGTGTG	1457

RESULT 6
US-09-476-242-7

```

1  GENERAL INFORMATION:
2  APPLICANT: BARNETT, Susan
3  APPLICANT: HARTOG, Karin
4  APPLICANT: MARTIN, Eric
5  TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
6  FILE REFERENCE: 1605.002
7  CURRENT APPLICATION NUMBER: US/09/476,242
8  CURRENT FILING DATE: 1999-12-30
9  NUMBER OF SEQ ID NOS: 26
10 SOFTWARE: PatentIn Ver.. 2.0
11 SEQ. ID NO. 7
12 LENGTH: 2334
13 TYPE: DNA
14 ORGANISM: Artificial Sequence
15 FEATURE:
16 OTHER INFORMATION: Description of Artificial Sequence: Leu122-Ser199
17 US-09-476-242-7
18
19 Query Match 2.9%; Score 69.8; DB 10; Length 2334;
20 Best Local Similarity 44.3%; Pred. No. 46-09;
21 Matches 422; Conservative 0; Mismatches 522; Indels 9; Gaps 3

```

Oy	462	ACCTGGGACTTGGGCGCGGCGCCCAACACCCCAAGACAGATGGCGAGGGCTCCGCGGC	521
Db	514	ACCCACGGCATCCGCCCGGTGTGAGACACCAGCTGTCTGTAAAGGACGCTTGGCCGAG	573
Oy	522	CTCGGCTTTGACGAGGTGTTTGACACGGCTGTTTGGCGCCGACCTGACCATCATGAGAG	581
Db	574	GAGGGCGGTGTGATCCGACGCGAGAACTTCAACCGACAACGCCAAGACATCATCTGTGAC	633
Oy	582	GGCAGCGGCTGTGCAACCGCTCACCGACACTGTGAGAGGCCACCCGATCTCCAGCAG	641
Db	634	CTGAAGGAGACGTGGAGATCAACTGCAACCGCCCCCAACAAACACCCGCAAGACATC	693
Oy	642	CCGCTGCCCATATGTTCAACGACTGTGCGCCCGGCTGATGTGCTATCTGGAGAAATCTTAC	701
Db	694	ACCATGGGCGCCCGCGCCCTTCTTACGCCACCGGCGACATATGTGGGCACATCCGCCAG	753
Oy	702	CCGGACTGATCCCTTACGTGAGCACTGCAGAGACCCCCAGATATGTGTGGCGGCATG	761
Db	754	GCCCACTGCACATCAGCGGCGAGAAATGGAAACAACCTTGAAAGATCTGTGACCAAG	813
Oy	762	GTCAGTCTTACTTACGCGGAAAAAGAGGCAATCGCCCAAGAGACATGTCATGTGTTC	821
Db	814	CTGCAGGCGCCAGTTCTGGCAACAGACCAATCGTGTTCAGACAGACGAGCGCGCCACCC	873
Oy	822	ATCATGCCCTGAGCGCGCAAGCAGTGGAGGCTGACCGGCACTGGTTCGTGTGACGCG	881
Db	874	GAGATCTGATGACACAGCTTCAACTGCGCGCGCGAGTTCTTACTGCAACAGCACCCAG	933
Oy	882	GACCCCACTTCCGCGCAGCTGAGCCACGTCATCACACCCTGTGAGACTGGCAACATCTTC	941
Db	934	CTGTTCAACAGACCTGGAAACAACACATCGGCCCAACACCAACGCGACCATTCAC	993
Oy	942	AAGAGCGGGGATTAACCTGGCGGAGCTGCGCGAGGGGAGTGGACATTCAAATGGGC	1001
Db	994	CTGCGCTTCCGATCAACGATCATCAACCCCTGGACGAGAGTGGCGAAGGCCATGTAC	1053
Oy	1002	GTGGGCTTGGGGCGCGGC---GTGCTGTTCCGACACACCGCGCGTGTCAATGAGCGGGG	1055
Db	1054	GCCCCCCCCATCCGCGGCACAGTCCGCTGACAGACACATACCGGCTCTGTGAC	1111
Oy	1059	CTGCGCACGCGCTATGAGCTGTTTCAACGGGACGCGCGCTGCGGCTGAGCTTGAGCGAG	1111
Db	1114	CGGACGCGCGGCAAGAGATCAAGCAACACACAGATCTTCCGCCCGCGCGGCGGCAC	1177
Oy	1119	GTCGCGGGCATGACGCAATCAAGAGACCAATCACCATGT--GCCCGCGCCGAG	1177
Db	1174	ATGCGCGCAACCTGCGGACGCGAGCTGTAAAGTACAAAGTGTGAAGATCGAGCCCCG	1233
Oy	1176	TCCAGTTTGAAGACCTGTGAAGACCGCGCGCGCGCGCGCGCGCGAGAGCCCGCGCGAC	1233
Db	1234	GAGGTGGCCCCCAACGAAGGCCGCGCGGTGTGACGCGGAGAAAGCGCCGCTGTACC	1297
Oy	1236	GGCACCCCCCGGCGCTGGCTGTGGAGACGGCGGCGGCTTACACAGCAGAGAGGAGAG	1299
Db	1294	CTGGGCGGCATTTCTTCTGGGCTTCTGTGGCGCGCGCGGAGAGCACAATGGCGCCCGGAC	1355
Oy	1296	GCGGCACTACACTGGCGGTGGCGGTGGCCAAAGGCTGGGCAACGCCAAGAGCTGATC	1355
Db	1354	CTGACCTTCAGACGTGTCA---GGCGCGCACGCTGTGAGCGGCGCATGTTGACAGCAGAAC	1411
Oy	1356	ACCAAGATGACAGCGCGGCGGACGGCCAAATACGATTTGTGAGATCATGTGCTTG	1408
Db	1411	AACCTGTGCGGCGCATGAGGCGCCAGACGCACTGTGTCACTACCTGATCGCTGTG	1463

```

:   APPLICANT: MARTIN, Eric
:   TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
:   FILE REFERENCE: 1605.002
:   CURRENT APPLICATION NUMBER: US/09/476,242
:   CURRENT FILING DATE: 1999-12-30
:   NUMBER OF SEQ ID NOS: 26
:   SOFTWARE: PatentIn Ver. 2.0
:   SEQ ID NO 25
:   LENGTH: 2358
:   TYPE: DNA
:   ORGANISM: Artificial Sequence
:   FEATURE:
:   OTHER INFORMATION: Description of Artificial Sequence: Val127-Asn195
US-09-476-242-25

```


Db 1324 CTGGGCGCATGTTCTGTGGGCTTCTGTGGGCGCGCGGCGACCAATGGGCGCGCGACG 1383
QY 1296 GGGGGCATCAGTCTGGGCGCGGTGGGCGCAAGCGGCTGGGCAAGCCCAAGAGTGTATC 1355
Db 1384 CTACCTGTGACCGTGTCA--GGCGCGCGCGAGTCTGTAGCGGCTGTCTGTGACAGCAGGAG 1440
QY 1356 ACCAAGATGAGCGCGCGGCGGCGCAAGTACGACTTTGTGAGATCATGGCCTG 1408
Db 1441 AACCTGTGGCGCGCATCGAGCGCGACGACGACCTGTCTGAGCTGACCGTGTG 1493

RESULT 8

US-09-808-880-1/c
; Sequence 1, Application US/09808880
; Publication No. US20030027287A1
; GENERAL INFORMATION:
; APPLICANT: Belach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/808,880
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US/09/428,517
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/120,254
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/106,100
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 50937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant DNA
US-09-808-880-1

Query March 2.7%: Score 64.6; DB 9; Length 50937;
Best Local Similarity 45.6%; Pred. No. 2.5e-07;
Matches 376; Conservative 0; Mismatches 419; Indels 30; Gaps 3;
QY 857 CCGCGACTGTTCTGTGTGACGCGGACCCGCGCGCAAGTGTGACCAAGTGTATC 916
Db 3774 CCGCGCTCGGGCGGAGGTTCCTCCGCGGTGGGGCGGTCTCTTACGACGCGACCAT 3715
QY 917 CACCGTGAAGTGGGCAATCTTTCAGAGAGCGGCGATCACTGGCGGAGTGGCCGA 976
Db 3714 CACCGTCCGGTCAACCGGACGATCGCGGCGCGCGCGGCGGCAAGCGCTA 3655
QY 977 GGGGAGTGGGCAATCAATGGGCGTGGGCGCGGCGCGGCGTGTGTGGGACGCGCT 1036
Db 3654 GGTGCAACTCGAGCGCGCGAGAGCGCTTGGCGGTGGCGCTGTGAGCGGAGCTAG 3595
QY 1037 CGGCGGTGTATGAGGCGGCGCTGCGGCGCGGCGCTATGAGTGTGTACGGGCGCGCT 1096
Db 3594 CAGCGCGCGCATGGCGGTGGCGCGGCGGAGCAAGCTGCGCATACCGGCGGCGCT 3535
QY 1097 GCGCGCGCTGAGCTGAGCGAGTGGCGGCGGCGATGACCGGATCAAGAGAGCAACATCAG 1156
Db 3534 GCGCAAGTGGCGCATGAGCTGTGAGCTGACGTGGTCACTCGCGAGGCGCGCGAGAGG 3475
QY 1157 CATGTGCGCGCGCGCGGCTCCAAATTGAGAGTGTCTGAAGCAGCGCGCGCGCGCG 1216
Db 3474 GATGACCGGG--CGGTGTGATGCGCGGAGCGGAGAGCGGAGTGTCTTGTGCGTA 3418
QY 1217 CGCGGAGCGCGCGCGGCGGAGCGGCGCGCGCGGCGGCTGTGGAGCGCGGCGCGGCTT 1276
Db 3417 CGAGAGCGGCTCGAGAGAGCGGAGCGGCGGAGCGGAGCTGTCTCGCGCGCGCGGAGGA 3358
QY 1277 CACGAGCGAGGAGCGGAGGCGGCGGATCACTGGCGGTGGCGCGCA----- 1327

Db 3357 CTGATCAGCGCGGTGGCGCGGTACCAACGAGACCTGGTGGTGTCTCAACAGCGGCTC 3298
QY 1328 -----CGGCTGGGCGMACCGCAAGGTGTATCAACAAAGTACGAGCGG 1372
Db 3297 CTCCTCAACATGCCCTGGCTGAGGAGACCGCGCGGTGTCTACCATGTGTACCGCGG 3238
QY 1373 CGAGGCGAAGTACGACTTGTG--TGAATATATGCTCTGCCCCGGGCTGTGTGGGCGG 1429
Db 3237 CGAGCGCGCGCGCGAGGCGAGCACCGGCTGTGTGGCGAGCGCGCAAGCGAGAGCGG 3178
QY 1430 CGCGCGCGAGCGCGGCTCCACCGGAGCAAGCCATCAAGCAAGCGGCGAGCGGCGCTGT 1489
Db 3177 GCTGACCCAGACTTTCGCGCGGAGAGAGGAGGAGAGCGGCTTGGCGGAGAGCGGCGG 3118
QY 1490 CAACCTGAGAGAAAGTCCAGCTCGCGCGCGGAGGAGGAGAAACCGCTCATCGGAGCT 1549
Db 3117 CTACCGGAGTGTGATGACCACTGACTACGAGGAGCTACTCGGCTACCGCTG 3058
QY 1550 GTACGACAGTACTCTGGAGAGCGGCTGGGCGGAGCGGAGCGGAGCTGTGACAGCCCA 1609
Db 3057 GTACGAAACGAGGCGGTCTCAGCGCGCTGTCTTGTGGCGAGCGGAGTGTGTCTACACTC 2998
QY 1610 CTACGTCGCGCGGCGGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1654
Db 2997 CTTCGACTACCGGAGCTGAGGTGACGCGGAGCGGCGGAGCGGCGG 2953

RESULT 9

US-10-171-311-51
; Sequence 51, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kametkar, Shubhangi
; APPLICANT: Giat, Karen
; APPLICANT: Ganavarapu, Manjula
; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 15952
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-171-311-51

Query March 2.7%: Score 63.8; DB 9; Length 15952;
Best Local Similarity 49.8%; Pred. No. 3e-07;
Matches 161; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
QY 1142 GGAAGCAATACATCAATGTGCGCGCGCGGTCCAAATTGAGAGCTGCTGAAGCA 1201
Db 8410 GAGAGCGGCGACCGGCTTGTATGACCCCGGTGGCAACTGAGCTGTGGTGGAGA 8469
QY 1202 CCGGCGCGCGCGCGCGGCGGAGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1261
Db 8470 GAGCGTGGCGCGCGGCTGTGGGCGGCGGAGATTCAGAGAGAGCTGTGTGGCGGAGG 8529

QY 1239 ACCCGGCGGCTGCTGGAGCGCGCGGCTTACGAGCGAGGAGCG 1298
 DB 1291 GGGCCATGTTCTGGGCTTCTGGGCGCGCGGAGCACATGCGCCGCGAGCTTG 1350
 QY 1299 GGCATCACACTGCGCGCTGGCGGCGGAGCGGCTGGGAGCGGCAAGAGCTGATCACC 1358
 DB 1351 ACCCTGACCGTGA---GGCGCGCAGCTGCTAGCGGCGATCGTGAGAGCAAGAAAC 1407
 QY 1359 AAGATGAGCGCGCGGAGCGGAGCGGAGTACGACTTTGTGAGATCATGCGCTG 1408
 DB 1408 CTGCTGGCGGCGCATCGAGGCGGAGCGAGCAGCCTGCTGAGCTGACCGTGTG 1457

RESULT 12

US-09-476-242-10
 ; Sequence 10, Application US/09476242
 ; Patent No. US20020146683A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BARNETT, Susan
 ; APPLICANT: HARTOG, Karin
 ; APPLICANT: MARTIN, Eric
 ; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
 ; FILE REFERENCE: 1605.002
 ; CURRENT APPLICATION NUMBER: US/09/476,242
 ; CURRENT FILING DATE: 1999-12-30
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 10
 ; LENGTH: 2541
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Arg426-Gly431
 US-09-476-242-10

Query Match 2.7%; Score 63.6; DB 10; Length 2541;
 Best Local Similarity 43.9%; Pred. No. 2e-07;

Matches 417; Conservative 0; Mismatches 524; Indels 9; Gaps 3;

QY 462 ACCCTGGGCTGCGCGCGCGCGGCGGAGCGGCGGCTTCCGCGC 521
 DB 733 ACCGACGGCATCCGCCGCTGTGTAGAGCCGAGCTGCTGAAAGCGAGCTTGCGGAG 792
 QY 522 CTCGGCTTTGACGAGGTGTTTGACACGCTGTTTGCGCGCGGAGCTGACCATATGAGAGAG 581
 DB 793 GAGGCGGTGTGATCCGACGAGAGAACTTCAACGACAAAGCGGAGCATCATGTCGAG 852
 QY 582 GCGAGCGAGCTGCTGCAACCGCTCAGGAGCACTGAGGCGGCGGAGCTCGAGAG 641
 DB 853 CTGAAGGAGAGCGTGGAGATCACTGACCGCGCGGCGGAGCAACAGAGAGATC 912
 QY 642 CCGGCTCCCATGTTTACCAAGCTGTGCGCGCGCTGATCGCTATGCTGAGAAATCTTAC 701
 DB 913 ACATCGGCGCGCGCGCGGCGCTTCTAGCGCAACGCGGAGCATCATGCGGAGATCCGCGAG 972
 QY 702 CCGGAGCTGATCCCTAGTGAAGAGCTGCAAGAGCGGCGGAGATGATGCTGCGCGCATG 761
 DB 973 GCGGACCTGCAACATCAAGCGGAGAGAGTGAACAAACCTGAAAGCAGATGTAACAAG 1032
 QY 762 GTCAAGTCTTACCTAGCGGAGAAAGAGGAGCATGCGGCGGAGAGAGATGATGATGCTG 821
 DB 1033 CTGACGCGCGCACTT---CGGCAACAGACATCTGTTTCAAGAGAGAGAGCGGCGGAGC 1089
 QY 822 ATCATGCGCTTGCACGCGAGAGAGTCTGAGAGCTGACCGGAGCTGCTGTGTGAGAGCGC 881
 DB 1090 CCGGAGATGTGATGACAGCTTCAATGCGGCGGCGAGTTCTTCTAATGAGAGAGCACC 1149
 QY 882 GACCCCACTGCGCGGAGCTGAGCAAGCTCATACACCTGAGAGCTGGGCAATCTTTC 941
 DB 1150 CAGCTGTTCAACAGCAGCTGGAACAAACATCGGCGGCGGAGCAACAGAGAGAGCAGATC 1209
 QY 942 AAGAGCGGCGGAGCTGAGCTGCGAGCTGCGGAGGCGGAGTGGAGCAATCAATGAGGCGC 1001

DB 1210 ACCCTGCGGCGCGGATCAAGCAATCATCAACCGCGGCGGAGCGGAGCGCATGAGCGC 1269
 QY 1002 GTGGCTGCGGCGCGGCGGCTGTTTGGCAACACCGGCGGCTGATGAGAGCGGCGCTG 1061
 DB 1270 CCCCCATCGGCGGCGGAGATCCGCTGCGAGAGCAATCAACCGGCGGCTGCTGAGCGCGC 1329
 QY 1062 CGCAGCGCTATGAGACTGTTTCAAGGCGGCGGCGGCGGCGGCTGAGCTGAGCGAGGTG 1121
 DB 1330 GAGGCGGCGAGAGATCAAGCAACACCGAGATCTTCCGCGGCGGCGGCGGAGCATG 1389
 QY 1122 CGCGCATGAGCGGATCAAGAGAGCAACATCAACATG---GCGGCGGCGGAGTCC 1178
 DB 1390 CGGCAACCTGCGGAGAGAGTGTACAAAGTGTGAGATGAGAGCGGCGGCGGCGC 1449
 QY 1179 AAGTTAGAGAGCTGTGAAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 1238
 DB 1450 GTGGCGGCGGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1509
 QY 1239 ACCCGGCGGCGGCTGCGGAGCGGCGGCGGCGGCGGCTTCAACGAGAGAGCGGAGCGC 1298
 DB 1510 GCGGCGATGTTCTGGGCTTCTGGGCGGCGGCGGCGGAGAGCATGAGCGGCGGCGGAGCTG 1569
 QY 1299 GGCATCACACTGCGGCTGCGCGCTGCGCAACGCGGCTGGCGGAGCGGAGAGCTATCAC 1358
 DB 1570 ACCCTGACCGTGA---GGCGCGCAGCTGAGCGGCGGCTGAGAGAGAGAGCAAC 1626
 QY 1359 AAGATGAGAGCGGCGGAGAGCGGAGTACGACTTTGTGAGATCATGCGCTG 1408
 DB 1627 CTGCTGGCGGCGCATCGAGGCGGAGGAGCAGCAGCCTGCTGAGCTGACCGTGTG 1676

RESULT 13

US-09-476-242-22
 ; Sequence 22, Application US/09476242
 ; Patent No. US20020146683A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BARNETT, Susan
 ; APPLICANT: HARTOG, Karin
 ; APPLICANT: MARTIN, Eric
 ; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
 ; FILE REFERENCE: 1605.002
 ; CURRENT APPLICATION NUMBER: US/09/476,242
 ; CURRENT FILING DATE: 1999-12-30
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 22
 ; LENGTH: 2298
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Val1120-11e201;
 US-09-476-242-22

Query Match 2.6%; Score 62.6; DB 10; Length 2298;
 Best Local Similarity 44.5%; Pred. No. 3.7e-07;

Matches 426; Conservative 0; Mismatches 519; Indels 12; Gaps 4;

QY 458 CGAGACCTGCGGCTGCGCGCGGCGGCGGAGCGGCGGAGCGGAGCTGCGGAGGCGCTCG 517
 DB 483 CGTGAACACCGTGCAGTGCACCGGAGATCCCGCGGCTGTGAGAGCAACGATGCTGCT 542
 QY 518 CCGGCTCGGCTTGAAGAGGTGTTGACACGCTGTTTGGCGGAGCTGACATCATGAGA 577
 DB 543 GAAAGGAGCTGCGGAGGAGGCGGCTGTGATCCGAGGAGAACTTCAACGAGCAAGC 602
 QY 578 GAGGCGGAGAGCTGCTGACCGGCTTCAACGAGCACTGAGAGCGGAGCGGAGCTTCCGA 637
 DB 603 CAAGACCATATGCTGAGCTGAGAGAGAGCGTGAAGATCA---CTGCAACCGGCGGCA 659
 QY 638 CAGGCGGCTGCAATGTTTCAACAGCTGCTGCGGCGGCTGATGCTATGCTGAGAAATC 697
 DB 660 CAACACACCGGAGAGAGATCACTGCGCGGCGGCGGCGGCTTCTTACGCGACCGGCGA 719

```

QY 698 TTAACCGGACCTGATCCCTACCTGAGCAGCTGCAAGAGCCCGCCAGATGATGCTGCGGC 757
DB 720 CATCATCGGCGACATCGCGCAGGCCCACTGCAACATCAGCGCGAGAAAGTGAACAACAC 779
QY 758 CATGTGCAAGTCTTACCTAGACGAAAAGAAAGGCAATCGCCCAAGACATGCTCATGCT 817
DB 780 CCGAAGCAGATGTGTACCAAGCTGACAGGCCAGTTGTGCAACAAAGCAATCTGTTCAC 839
QY 818 GTTCATCATGCTCTGACCGCGCAAGCAGTGCAGAGCTGACCGGCACTGTTCTGTGGA 877
DB 840 GCGAGACAGCGGCGGCGACCCCGAGATCGTATGCAACGTTCAATCTGCGCGCGAGTT 899
QY 878 CCGCGACCCCACTCGGCGCAGCTGGAACAAGTATCAACAAGTGTGAGCTGCGCAACAT 937
DB 900 CTTCCTCTGCAACAGACAGCAGCTGTTCACACAGCACTGGAACAACATCTGCGCCCA 959
QY 938 CTTCAGAGAGCGGCGGATCAACTGTGCGCAGAGCTGCGCGAGGCGAGTGGGACATCCAT 997
DB 960 CAACACCAACGGGACCATCACTGCTGCTGCGCGCATCAAGATCATTCGCGCGCGCAT 1019
QY 998 GGGCGTGGGCTCGGGCGCGCGCGC---GTGCTGTTGCGGACCAAGCGCGGCTGTCATGAGGC 1054
DB 1020 GTACGCGCCCGCCATCGCGCGCGCAGATCGCTGACAGACAAATCAACCGGCTGCTGCT 1079
QY 1055 GGGCGTGGCGACCGGCTATGAGCTGTTCAAGGCGACCGCGCTGCGGCTGAGCCTGAG 1114
DB 1080 GACCGCGACGGGCGGAGGATCAAGCAACACACAGATCTTCGCGCGCGCGCGCG 1139
QY 1115 CGAGGTGCGCGGATGAGCGGCTCAAGAGACCAACATACATGCT---GCCCGCGC 1171
DB 1140 CGACATGCGCGCAACATGCGCGCAGCGCTGTACAGTACAGTGTGTGAAGATCGAGCC 1199
QY 1172 CGGCTCAAGTTTGAGAGCTGCTGAAGACGCGCGCGCGCGCGCGCGCGCGCG 1231
DB 1200 CTTGGGCTGTGCGCCCAACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1259
QY 1232 GCAAGGACCGCGCGCGCGCTGCGCTGGAAGCGCGCGCGCGCGCTTCAACAGCGAGCG 1291
DB 1260 GACCTGTGGGCGCATGTTCTTGGGCTTCTGCGCGCGCGCGCGCGCGCGCGCGCG 1319
QY 1292 CAGGGCGGCGATCACTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1351
DB 1320 CAGCCTGACCTGACCTGTGA---GGCGCGCAGCTGTGAGCGGCGATCTGTGAGCAGCA 1376
QY 1352 GATCACCAAGATGACGCGCGCGCGCGAGGCGAGTACGATTTGTGAGATCATGCGCTG 1408
DB 1377 GAACAACTGTGCGCGCGCATGAGGCGCGAGCAGCACTGTGCTGAGCTGACCTGTG 1433

```

RESULT 14

US-09-476-242-23
Sequence 23, Application US/09476242

Patent No. US2002014683A1

GENERAL INFORMATION:

APPLICANT: BARNETT, Susan

APPLICANT: HARTOG, Karin

APPLICANT: MARTIN, Eric

TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES

FILE REFERENCE: 1605.002

CURRENT APPLICATION NUMBER: US/09/476,242

NUMBER OF SEQ ID NOS: 26

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 23

LENGTH: 2298

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:

US-09-476-242-23

Query Match 2.6%; Score 62.6; DB 10; Length 2298;
Best Local Similarity 44.5%; Pred. No. 3.7e-07;
Matches 426; Conservative 0; Mismatches 519; Indels 12; Gaps 4;

```

QY 458 CGAGACCTCGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 517
DB 483 CGTGAGCAGCTGTGACATGACCCAGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 542
QY 518 CCGGCTCGGCTTTGACGAGGTGTTGACAGCGCTGTTTGCGCGCGCGCGCGCGCGCGCG 577
DB 543 GAAAGCGACCTGTGCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 602
QY 578 GGAAGGCGACGAGCTGTGACCGCGCTCAACCGACAGCTGGAAGCGCGCGCGCGCGCG 637
DB 603 CAAGACATATCATGTGTGACGTGAGAGAGCGTGAATCAA---CTGACCGCGCGCGCA 659
QY 638 CGAGCGCGCTGCCATTTACCAAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 697
DB 660 CAACAAACCGCGAGAGCATCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 719
QY 698 TTAACCGGACCTGATCCCTACCTGAGCAGCTGCAAGAGCCCGCCAGATGATGCTGCGGC 757
DB 720 CATCATCGGCGACATCGCGCAGGCCCACTGCAACATCAGCGCGAGAAATGGAACAAC 779
QY 758 CATGTGCAAGTCTTACCTAGACGAAAAGAAAGGCAATCGCCCAAGACATGCTCATGCT 817
DB 780 CCGAAGCAGATGTGTACCAAGCTGACAGGCCAGTTGTGCAACAAAGCAATCTGTTCAC 839
QY 818 GTTCATCATGCTCTGACCGCGCAAGCAGTGCAGAGCTGACCGGCACTGTTCTGTGGA 877
DB 840 GCGAGACAGCGGCGGAGCGCGCGCGAGTCTGTATGCAACATTTCACTGTGCGCGCGCA 899
QY 878 CCGCGACCCCACTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 937
DB 900 CTTCCTCTGCAACAGACAGCAGCTGTTCAACAGCACTGGAACACACCATGCGCGCGCA 959
QY 938 CTTCAGAGAGCGCGGCGATCACTGCGCGAGCTGCCGAGGCGAGTGGGACATCCAT 997
DB 960 CAACACCAACGGGACCATCACTGCTGCTGCGCGCATCAAGATCATTCGCGCGCGCAT 1019
QY 998 GGGCGTGGGCTCGGGCGCGCGCGC---GTGCTGTTGCGGACCAAGCGCGGCTGTCATGAGGC 1054
DB 1020 GTACGCGCCCGCCATCGCGCGCGCAGATCGCTGACAGACAAATCAACCGGCTGCTGCT 1079
QY 1055 GGGCGTGGCGACCGGCTATGAGCTGTTCAAGGCGACCGCGCGCGCGCGCGCGCGCG 1114
DB 1080 GACCGCGACGGGCGGAGGATGAGCAACACACAGATTTTCGCGCGCGCGCGCGCG 1139
QY 1115 CGAGGTGCGCGGATGAGCGGATCAAGAGACCAACATACCATGCT---GCCCGCGC 1171
DB 1140 CGACATGCGCGCAACATGCGCGCAGCGCTGTACAGTGTGTGAAGATCGAGCC 1199
QY 1172 CGGCTCAAGTTTGAGAGCTGCTGAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1231
DB 1200 CTTGGGCTGTGCGCCCAACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1259
QY 1232 GCAAGGACCGCGCGCGCGCTGCGCGCTGGAAGCGCGCGCGCGCGCTTCAACAGCGAGCG 1291
DB 1260 GACCTGTGGGCGCATGTTCTTGGGCTTCTGCGCGCGCGCGCGCGCGCGCGCGCG 1319
QY 1292 CAGGGCGGCGATCACTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1351
DB 1320 CAGCCTGACCTGACCTGTGA---GGCGCGCAGCTGTGAGCGGCGATCTGTGAGCAGCA 1376
QY 1352 GATCACCAAGATGACGCGCGCGCGCGAGGCGAGTACGATTTGTGAGATCATGCGCTG 1408
DB 1377 GAACAACTGTGCGCGCGCATGAGGCGCGAGCAGCACTGTGCTGAGCTGACCTGTG 1433

```

RESULT 15

US-09-476-242-24
Sequence 24, Application US/09476242

Patent No. US2002014683A1

GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: HARTOG, Karin
APPLICANT: MARTIN, Eric
TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
FILE REFERENCE: 1605.002
CURRENT APPLICATION NUMBER: US/09/476,242
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 2298
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: VAI120-Thr202;
OTHER INFORMATION: 11e424-A1a433
US-09-476-242-24

Query Match 2.6%; Score 62.6; DB 10; Length 2298;
Best Local Similarity 44.5%; Pred. No. 3.7e-07;

Matches 426; Conservative 0; Mismatches 519; Indels 12; Gaps 4;

QY 458 CGAGACCTGGGCTGGCGCGCGCGCCACACCCCAAGCAGCTGGCGGAGGCGCTCCG 517
DB 483 CGTAGACCGGTGAGTGCACCGGCGATCCGCCGTGTAGACACCCAGCTGTGCT 542
QY 518 CCGCCTGGCTTTGACGAGGTGTTGACAGCTGTTGGCGCCGACCTGACCATCATGA 577
DB 543 GAAGGACGCTGGCCGAGAGGCGCTGTGTATCCGACGAGAACTTCACTCGACACG 602
QY 578 GGAGGAGGAGCTGTGTGACCGCTTCAACGACACCTGGAGGCCACCGCATCTCGA 637
DB 603 CAAGACCATCATCTGTGAGCTGAAGAGAGGTGAGATCAA--CTGCAACCGCCCA 659
QY 638 CGAGCCCGCTGCGCATGTTTCAACAGCTGTGCGCGCTGAGTGGCTATGCTGAGAAATC 697
DB 660 CAACAAACACCGGCAAGAGCATACCATCGGCCCGCGCGCTTTTCAACGACCGGCGA 719
QY 698 TTACCCGAGCTGATCCCTTACGTGAGCAGCTGCAGAGCCCGCAGATGATGTGGCG 757
DB 720 CATCATGGGCGACATCCCGGAGGCCCACTGCAATCAGCGGCGAGAAAGTGAACAC 779
QY 758 CATGTCAGTCTTACTTACCTGAGGAAAGAGGSCATCGCCCAAGGACATGTGTATGT 817
DB 780 CCTGAGCAGATCTGTGACCAAGCTGAGGCCCAAGTTCGCAACAGACCATCTGTCTCA 839
QY 818 GTCCATCATGTGCTGTGACGCGCCAGCACTGGAGGCTGACCGGCACTGTGTGTGA 877
DB 840 GCAGAGCAGCGCGCGGCGACCCCGAGATCTGTATGACAGCTTCACTGCGCGGAGTT 899
QY 878 CGCCGACCCCACTGCGCAGCTGSAACAGTCATCAACAACGCTGAGCTGGCAACAT 937
DB 900 CTTTACTGCAACAGACCACTGTTTCAACAGCACTGGAACAAACATCGCCCCCA 959
QY 938 CTTCAAGGAGCGCGCATCAACTGTGCGAGCTGCGGAGGCGAAGTGGACATCCAT 997
DB 960 CAACACCAACGCGCACCATCACTGCTGCGCATCAAGCAGATCAATCGGCGGCGCAT 1019
QY 998 GGGCGTGGGCTCGGGCGCGCGG---GTGCTGTTGGGACCAACCGGCGGTGATGAGAGC 1054
DB 1020 GTACGCGCGCGCATCGCGCGCAGATCGCTGAGCAGCAACATCAACGCGCTGTGCT 1079
QY 1055 GAGCGTGGCGACGCGCTTATGAGCTGTTACGCGGACGCGCGCTGCGCGCTGAGCTGAG 1114
DB 1080 GACCGCGACCGCGCGCAAGAGATCAGCAACACGAGATCTTCCGCCCCGGGCGG 1139
QY 1115 CGAGGTGCGCGCATGACGCGCATCAAGAGACCAACATCAACATGT--GCCGCGCC 1171
DB 1140 CGACATGCGCGCAACTGCGCGCAGCGAGCTGTACAAAGTCAAGGTGTGAAGATGAGCC 1199
QY 1172 CGGCTCAAGTTTGAAGAGCTGTGAACACCGCGCGCGCGCGCGGCGGCGCGCC 1231

DB 1200 CCTGGCGCGTGGCCCCCACCAGAGGCCAGCGCGGTGTGACGCGAGAAACGCGCGCT 1259
QY 1232 GCACGCAACCCCGCGCGCTGTGCTGGAACGCGCGCGGCTTCAACAGCAGAGACG 1291
DB 1260 GACCTGGCGCGCATGTTCTGTGCTTCTGTGCGCGCGCGCAGACCATGCGCGCG 1319
QY 1292 CAGGGCGCGCATCAACACTGCGCGTGGCCCAACGGGCTGGGCAACGCAAGAGCT 1351
DB 1320 CAGCTGACCTGACCTGCA--GGCCCGCAGCTGCTGACGCGCATCTGCAAGCA 1376
QY 1352 GATCAACCAAGATGACGCGCGCGGAGCCAGTACGACTTTGTGAGATCATGCGCTG 1408
DB 1377 GAACAACTGTGCGCGCGCATGAGCGCCAGCAGCACTGTGACAGTGAACGTGTG 1433

Search completed: June 8, 2003, 14:36:16
Job time : 383 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 09:51:01 ; Search time 3028 Seconds
(without alignments)
12831.233 Million cell updates/sec

Title: AJ012098

Perfect score: 2399
Sequence: 1 ATCTTACATGACACACAA.....TGTAAGCAGCACTTCTGAG 2399

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gse:*
18: em_gse_hum:*
19: em_gse_inv:*
20: em_gse_pln:*
21: em_gse_vrt:*
22: em_gse_fun:*
23: em_gse_mam:*
24: em_gse_mus:*
25: em_gse_oher:*
26: em_gse_pro:*
27: em_gse_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	649	27.1	657	13	BI717904 1031022F1
2	637.4	26.6	648	13	BI532215 1024119G1
3	637	26.6	644	13	BM003317 1031109G0
4	614	25.6	622	13	BI716102 1031001E0
5	612.8	25.5	617	13	BI999849 1031079E0
6	596.4	24.9	601	12	BE724427 894076B10

7	556.4	23.2	558	10	AV389547
8	545	22.7	602	12	BG48385
9	543.8	22.7	599	13	BI716925
10	539.8	22.5	621	12	BG48384
11	536.4	22.4	538	10	AV396217
12	499	20.8	499	10	AV396237
13	498	20.8	498	10	AV386674
14	477.6	19.9	550	13	BI717903
15	476	19.8	631	13	BI532214
16	474.4	19.8	550	10	BE352293
17	471.2	19.6	541	10	BE351941
18	468	19.5	549	12	BE724426
19	464.8	19.4	490	10	BE352300
20	456.8	19.0	481	12	BE726557
21	455.4	19.0	717	12	BG48383
22	453.4	18.9	463	10	AV390537
23	446.2	18.6	456	13	BM003316
24	442.4	18.4	456	10	AV394544
25	422.4	17.6	509	13	BI994221
26	421	17.5	458	10	BE237741
27	414	17.3	415	10	AV639539
28	397.8	16.6	762	12	BF661236
29	384	16.0	385	12	BE725875
30	377	15.7	377	10	AV641632
31	374.8	15.6	393	10	AV389590
32	366.2	15.3	658	13	BI726441
33	362.8	15.1	436	10	AV634899
34	354	14.8	354	10	AV631034
35	352.8	14.7	641	13	BI875012
36	347.4	14.5	362	10	AW721316
37	333.2	13.9	585	10	AV389530
38	332	13.8	353	13	BI994585
39	312.8	13.0	482	10	AV395481
40	301	12.5	549	10	BE337478
41	292.4	12.2	308	13	BI725565
42	275.4	11.5	326	12	BE761208
43	268.4	11.2	573	13	BI529458
44	251.6	10.5	504	10	BE352294
45	235.2	9.8	505	10	BE452899

ALIGNMENTS

RESULT 1
LOCUS BI717904 657 bp mRNA linear EST 19-SEP-2001
DEFINITION 1031022F12.Y1 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BI717904
VERSION BI717904.1 GI:15693599

KEYWORDS

SOURCE

ORGANISM

REFERENCE
AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermost, J. P., Shrager, J., Sillitow, C. and Stern, D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model.
JOURNAL Unpublished (2001)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chause@duke.edu
Location/Qualifiers
1. 657
/organism="Chlamydomonas reinhardtii"

FEATURES

source

/stratn="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, Stress II (normalized
), lambda Zap II"
 /note="vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
 XhoI; Stress condition II library, constructed by John
 Davies and Jeffrey McDermott, combines cDNAs from CC-1690
 cells grown to mid-log phase in TAP (NH4+ - containing)
 and shifted to TAP - NO3- (24hrs): H2 production
 conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
 Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
 sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
 POLYA mRNA was purified from each sample, pooled and cDNA
 synthesized. The cDNA was directionally cloned into lambda
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
 sites. pBluescript II SK- plasmids were excised from the
 lambda Zap clones by superinfection with Exbaesist
 (Stratagene) phage. The library was normalized using
 method 4 described in Bonaldi et al., (1996) Genome
 Research 6: 791-806."

```

Query Match      27.1%; Score 649; DB 13; Length 657;
Best Local Similarity 100.0%; Pred. NO. 2e-12;
Matches 649; Conservative 0; Mismatches .0; Indels 0; Gaps 0.

```

QY	19	AACACTCTCGGAGGGAAGCTAAGCTCAAAACCTGGAACCTTTTCCAAAGTTTACACCC	78
Db	9	AACACTCTCGGAGGGAAGTAAAGCTCAAAACCTGGAACCTTTTCCAAAGTTTACACCC	68
QY	79	AATTGGAAGCGCGCTCCAAAGTCGCTCGATTGCTCTTCATGCAACCACTTAATTCT	138
Db	69	AATTGGAAGCGCGCTCCAAAGTCGCTCGATTGCTCTTCATGCAACCACTTAATTCT	128
QY	139	AATATCTGTAAGCGCACAAGATGTGGGCGCTGTGTGTAAGCCCTGCGGGCGGTGTCTA	198
Db	139	AATATCTGTAAGCGCACAAGATGTGGGCGCTGTGTGTAAGCCCTGCGGGCGGTGTCTA	188
QY	199	TTGCGGCGAGCTCTTGCAAGGCGGCGAGGTGAGCCCGCGCGCTCCGCTCGAGCGAGA	258
Db	189	TTGCGGCGAGCTCTTGCAAGGCGGCGAGGTGAGCCCGCGCGCTCCGCTCGAGCGAGA	248
QY	259	CCGTGCTGTACCCCTTGGCAACTTGAAGGCGCGCGACCGCGCTAAGCAACGTGCTT	318
Db	249	CCGTGCTGTACCCCTTGGCAACTTGAAGGCGCGCGACCGCGCTAAGCAACGTGCTT	308
QY	319	GGGCGGCTGCGCGACCCGCGTGGGAGGCGCCTTTGATGTCATGTCGAGAGGCGCTCGCG	378
Db	309	GGGCGGCTGCGCGACCCGCGTGGGAGGCGCCTTTGATGTCATGTCGAGAGGCGCTCGCG	368
QY	379	AGCTTGCCAAAGCCCAAGAGAGAACCCCAAGCGAGCAAGCAAGCTCTGCGTAGGTGCTCCGG	438
Db	369	AGCTTGCCAAAGCCCAAGAGAGAACCCCAAGCGAGCAAGCAAGCTCTGCGTAGGTGCTCCGG	428
QY	439	CCGTTGCTGTCCGTATTGCTCCGAGAACCTTGAGCGCTGGCGCGGGCGGCACACCCCAAC	498
Db	429	CCGTTGCTGTCCGTATTGCTCCGAGAACCTTGAGCGCGCTGGCGCGGGCGGCACACCCCAAC	488
QY	499	AGCTGGCCGAGGGGCGCTCGGCGCGCTCGGCTTTGACAGAGGTATTGACAGCTGTTGGCG	558
Db	489	AGCTGGCCGAGGGGCGCTCGGCGCGCTCGGCTTTGACAGAGGTATTGACAGCTGTTGGCG	548
QY	559	CCGACTGACCAATCATGAGAGAGAGGAGAGCAAGCTGTGTGACCGGCTCACCGAGCACTGG	618
Db	549	CCGACTGACCAATCATGAGAGAGAGGAGAGCAAGCTGTGTGACCGGCTCACCGAGCACTGG	608
QY	619	AGGCCCCACCCGCACTCCGACGAGAGCGCTGCCCATGTTCACCAACTGCTG 667	
Db	609	AGGCCCCACCCGCACTCCGACGAGAGCGCTGCCCATGTTCACCAACTGCTG 657	

RESULT 2

B1532215 648 bp mRNA linear EST 29-AUG-2001

LOCUS 1024119G12.Y1 C. reinhardtii CC-1690, normalized, Lambda Zap II

DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION B1532215 GI:15372789

VERSION B1532215.1

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii.

ORGANISM Chlamydomonas reinhardtii.

REFERENCE Eukaryote; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 648)

AUTHORS Groseman,A., Chang,C.-W., Davies,J.J., Harris,E., Hauser,C., Lefebvre,
P., McDermott,J.P., Shresler,J., Siffow,C. and Stern,D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1024b
Unpublished (2001)
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel.: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

JOURNAL COMMENT Location/Organism

TITLE

FEATURES

source 1..648

```

SOURCE
1. .648
   /organism="Chlamydomonas reinhardtii"
   /strain="CC-1690 wild type mt+ 21gr"
   /db_xref="taxon:3055"
   /clone_lfb="C. reinhardtii CC-1690, normalised, lambda Zap
   II"
   /note="Vector: pbluescript II SK-, site 1: EcoRI; Site 2:
   XhoI; This library, constructed by John Davies and Jeffrey
   McDermott, combines cDNAs from CC-1690 cells grown to
   mid-log phase in TAP (acetate-containing) medium in the
   light, TAP medium in the dark, HS (minimal) medium in
   ambient levels of CO2 and HS medium bubbled with 5% CO2.
   PolyA mRNA was purified from each sample, pooled and cDNA
   synthesized. The cDNA was directionally cloned into lambda
   Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites
   pbluescript II SK- plasmids were excised from the lambda
   Zap clones by superinfection with Exsist (Stratagene)
   phase. The library was normalized using method 4 described
   in Bonaldo et al (1986) Genome Research 6: 791-806."

```

Query Match	Similarity	26.6%	Score 637.4	DB 13	Length 648
Best Local	Similarity	98.9%	Pred. No. 4.5e-119		
Matches	641	Conservative	0	Mismatches	7
				Indels	0
				Gaps	0
QY	17	CAAACTCTCTGCAAGGACACTAGCCTCAAAACCTCGAAACCTTTTTCAAACAGTTTACAC	76		
Db	1	CAAACTCTCTGCAAGGACACTAGCCTCAAAACCTCGAAACCTTTTTCAAACAGTTTACAC	60		
QY	77	CCAATTGGAGCGCGCTCCAAAGCTCGCTCCGTTGCTCTTCAATGCAACACTATTATT	136		
Db	61	CCAATTGGAGCGCGCTCCAAAGCTCGCTCCGTTGCTCTTCAATGCAACACTATTATT	120		
QY	137	CTAATATGCTAGACCGCAAAAGATGTGGGCGTGGTGAAAGCCTGCGCGCGCGTATC	196		
Db	121	CTAATATGCTAGACCGCAAAAGATGTGGGCGTGGTGAAAGCCTGCGCGCGCGTATC	180		
QY	197	TATTGCGGAGAGCTCTGCAAGGCGGCGAGGTCGCCCGCGCTCCGCTGCGACGAG	256		
Db	181	TATTGCGGAGAGCTCTGCAAGGCGGCGAGGTCGCCCGCGCTCCGCTGCGACGAG	240		
QY	257	CACCGTGGGTATGAGCCCTTTGCAACCTTGAAGGCGCGGCAAGCGCGCTTAGGCAAGTGGC	316		
Db	241	CACCGTGGGTATGAGCCCTTTGCAACCTTGAAGGCGCGGCAAGCGCGCTTAGGCAAGTGGC	300		
QY	317	TTGCGGAGCTCGGCAACCCGCTGCGGAGGCGCTTTGATGATATGTCAGACAGGCGCTGGC	376		

RESULT 2

Db 301 TTGGCGGCTGCGCGACCGCGCTGGAGGCGCTTTGAGTATGTCAGAGGCGCTCGC 360

QY 377 CGAGCTTGCCAAAGCCCGANAGACGACCCCA CGCGCAGA GCGTCTGCGTCAGAGTGGCTCC 436

Db 361 CGAGCTTGCCAAAGCCCGANAGACGACCCCA CGCGCAGA GCGTCTGCGTCAGAGTGGCTCC 420

QY 437 GGGCGTTGCTGTGCGTATTTGGCGAGAACCTGGGCGCTGGCGCGCGGGGGGCACACGCCCCAA 496

Db 421 GGGCGTTGCTGTGCGTATTTGGCGAGAACCTGGGCGCTGGCGCGGGGGGCACACGCCCCAA 480

QY 497 GCAGCTGCGCGAGGGGCGCTCCGCGCGCTCGGCTTTTGA CGAGGTGTTTGA CACGCTGTTTGG 556

Db 481 GCAGCTGCGCGAGGGGCGCTCCGCGCGCTCGGCTTTTGA CGAGGTGTTTGA CACGCTGTTTGG 540

QY 557 CGCGCACTGACCATCATGAGAGAGGGCAGCGAGCTGCTGCA CGCGCTTCA CGAGACACT 616

Db 541 CGCGCACTGACCATCATGAGAGAGGGCAGCGAGCTGCTGCA CGCGCTTCA CGAGACACT 600

QY 617 GGAGGCCACCGCGACTCGAGCGAGCGCGTGGCCATGTTTACACGACTG 664

Db 601 GGAGGCCACCGCGACTCGAGCGAGCGCGTGGCCATGTTTACACGACTG 648

LOCUS 644 bp mRNA linear EST 25-OCT-2001
 EM030317 1031109G06.Y1 C. reinhardtii CC-1690, Stress II (normalized),
 Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION	BM003317
VERSION	BM003317.1
KEYWORDS	GI:16438097
SOURCE	EST.
ORGANISM	<i>Chlamydomonas reinhardtii</i>
	<i>Chlamydomonas reinhardtii</i>

REFERENCE
Eurytycea; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 644)

AUTHORS	TITLE
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shreager, J., Silflow, C. and Stern, D.	Analyses of the <i>Chlamydomonas reinhardtii</i> Genome: A Model,

JOURNAL
COMMENT
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished (2001)
Contact: Charles Hauser

DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES	Location/Qualifiers
source	1. .644

"organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt⁻ 21gr"
 /db xref="taxon:1055"
 /clone.lib="cc_reinhardt CC-1690, Stress II (normalized
), lambda Zap I"
 /nose=Vector: pBluescript II SK⁺, Site_1: EcoRI; Site_2:
 XhoI; Stress condition II library, constructed by John
 Davies and Jeffrey McDermott, combines cDNAs from CC-1690
 cells grown to mid-log phase in TAP (NH₄⁺ - containing)
 and shifted to TAP - NO₃⁻ (24hrs); H2 production
 conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
 Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
 sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
 PolyA mRNA was purified from each sample, pooled and cDNA
 synthesized. The cDNA was directionally cloned into lambda
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
 sites. pBluescript II SK⁺ plasmids were excised from the
 lambda Zap clones by superinfection with Exsist
 (Stratagene) phage. The library was normalized using
 method 4 described in Bonaldo et al., (1996) Genome
 research 6:791-806."

BASE COUNT	118 a	213 c	230 g	83 f
ORIGIN				

Query Match	26.6%	Score 637;	DB 13;	length 644;
Best Local Similarity	100.0%	Pred. No. 5.4e-119;		
Matches 637;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy 755 GGCCATGGTCAAGTCTACTTAGCGGAAAAGAAG6CATCGGCCCAAAGGACATGGTCAT 814
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8 GGCCATTGGTCAAGTCTACTTAGCGGAAAAGAAG6CATCGGCCCAAAGGACATGGTCAT 67

Qy	815	G	G	T	C	C	A	T	G	C	C	T	G	C	C	G	A	A	G	A	G	T	C	G	A	C	C	G	A	C	T	T	G	T	874
Db	68	G	G	T	C	C	A	T	G	C	C	T	G	C	C	G	A	A	G	A	G	T	C	G	A	C	C	G	A	C	T	T	G	T	127

[illegible][illegible]

995 AATGGCGCTGGGCTCGGGCGTGTCTTCGGACACACCGCGGTCTCATGAGGC 105
248 AATGGCGCTGGGCTCGGGCGTGTCTTCGGACACACCGCGGTCTCATGAGGC 307

QY 1055 GCGCTGAGCAGGCGCTATGAGCTGTTCA CGGCGACGCCGCTGCCGCGCTGAGCCTGAG 1111
|||
Db 308 GCGCGTCGCGCAGCGCGCTATGAGCTGTTCA CGGCGACGCCGCTGCCGCGCGCTGAGCCTGAG 367

QY 1115 CGAGGTGCGCGGCATGAGCGGCATCAAGGAGCAACATCACCATTGTCGCCGCCCGG 117

Db 368 CGAGGTGCGCGGCATGAGCGGCATCAAGGAGCAACATCACCATTGTCGCCGCCCGG 422

Qy 1175 GTCCAGTTTGAAGAGCTCTGAACACCGCCGCCGCCGCCGAGCCGCCGCA 1231
|||||
Db 428 GTCCAGTTTGAAGAGCTCTGAACACCGCCGCCGCCGCCGAGCCGCCGCA 487

Qy 1235 CGGCACCCCCGGGCGCTGGGACGGCGGCGGGCTTCAACGACGAGACGGCAG 129

Oy 1295 GGGCGGATCAACCTGCGGCTGGCCAAAGGCTGGGCAACGCCAAGCTGAT 1351
 |||||
 Db 548 GGGCGGATCAACCTGCGGCTGGCCAAAGGCTGGGCAACGCCAAGCTGAT 604
 |||||

0y 1355 CACCAAGATGCAGCGCGGAGCCCAAGTACGACTTT 1391
 |||||
 Db 608 CACCAAGATGCAGCGCGGAGGCCCAAGTACGACTTT 644
 |||||

RESULT 4
81716102

LOCUS	B716102	622 bp	mRNA	linear	EST 19-SEP-2000
DEFINITION	1031001E05.y1 C. reinhardtii CC-1690, Stress II (normalized),				
ACCESSION	Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.				
	B716103				

RECORDION	01/10/02
VERSION	B1716102.1
KEYWORDS	GI:15691797
SOURCE	EST.
ORGANISM	Chlamydomonas reinhardtii.

ORGANISM
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
1 (bases 1 to 622)
Greenberg A, Chava C, W. Dauter J, Haeberli P, Haeberli C, Tschopp
REFERENCE

TITLE Analyses of the *Chlamydomonas reinhardtii* Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Plants
AUTHORS Glushko, A., Chang, C.-M., Davies, J., Hall, B., Naud, C., P., McDermott, P., Shrager, J., Silflow, C., and Stern, D.

JOURNAL
vascular frames, project: 1031
unpublished (2001)
COMMENT
Contact: Charles Hauser
DCMB Box 91000

Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
Location/Qualifiers

FEATURES

source

1. 622
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mc+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
), Lambda Zap II"
/note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr; see Wells et al., (2000). Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. Bluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."

BASE COUNT 102 a 235 c 178 g 107 t
ORIGIN

Query Match 25.6%; Score 614; DB 13; Length 622;
Best Local Similarity 100.0%; Pred. No. 2,4e-114;

Matches 614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

17 CAAACACTCTGCGAGCGACTAGCTCAAAACCTTTTCCAGACGTTTAAAC 76
   |||||
9  CAAACACTCTGCGAGCGACTAGCTCAAAACCTTTTCCAGACGTTTAAAC 68
   |||||
77 CCAATTGGAGCGCGCTCCAGCTCGCTCGTCTCTCATGCGACCACTTATTT 136
   |||||
69 CCAATTGGAGCGCGCTCCAGCTCGCTCGTCTCTCATGCGACCACTTATTT 128
   |||||
137 CTAATATGTAAGCGCGACGATGTCGCGCTCTGTCTGAAGCTTGTGCGCGCTGTC 196
   |||||
129 CTAATATGTAAGCGCGACGATGTCGCGCTCTGTCTGAAGCTTGTGCGCGCTGTC 188
   |||||
197 TATTGCGCGAGCTCTCTGAGGCGCGCGAGCTGCGCGCTCGCTCGAGCGAG 256
   |||||
189 TATTGCGCGAGCTCTCTGAGGCGCGCGAGCTGCGCGCTCGCTCGAGCGAG 248
   |||||
257 CACGCTGCTGTAGCGCTTGCACACTTGAAGCGCGCGCGCTGAGCAAGCTGCG 316
   |||||
249 CACGCTGCTGTAGCGCTTGCACACTTGAAGCGCGCGCGCTGAGCAAGCTGCG 308
   |||||
317 TTGCGCGGCTGCGCGAGCTGCGAGGCGCTTGTGATGTCATGTCACAGCGCTGCG 376
   |||||
309 TTGCGCGGCTGCGCGAGCTGCGAGGCGCTTGTGATGTCATGTCACAGCGCTGCG 368
   |||||
377 CGAGCTTGCAGAGCGCGAGCGAGCGCGAGCGCGAGCGCGAGCGCGAGCTGCG 436
   |||||
369 CGAGCTTGCAGAGCGCGAGCGAGCGCGAGCGCGAGCGCGAGCGCGAGCTGCG 428
   |||||
437 GCGCGTGTGTGCTATTTGCGAGACCTTGGGCTTGGCGCGCGCGCGAGCGCG 496
   |||||
429 GCGCGTGTGTGCTATTTGCGAGACCTTGGGCTTGGCGCGCGCGCGAGCGCG 488
   |||||
497 GCGCGTGTGTGCTATTTGCGAGACCTTGGGCTTGGCGCGCGCGCGAGCGCG 556
   |||||
489 GCGCGTGTGTGCTATTTGCGAGACCTTGGGCTTGGCGCGCGCGCGAGCGCG 548
   |||||
557 GCGCGTGTGTGCTATTTGCGAGACCTTGGGCTTGGCGCGCGCGCGAGCGCG 616
   |||||

```

Db 549 GCGCGACTGACCATCATGAGAGGCGAGCGAGCTGCTGACCGGCTTACCGAGCACT 608
 Oy 617 GCGAGCGCGCGCGCG 630
 Db 609 GCGAGCGCGCGCGCG 622

RESULT 5
B1999849 617 bp mRNA linear EST 25-OCT-2001
LOCUS 1031079604.x2 C. reinhardtii CC-1690, Stress II (normalized),
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION B1999849
VERSION B1999849.1 GI:16434623

KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
Chlamydomonadaceae; Chlorophyta; Chlorophyceae; Volvocales;
Eukaryota; Viridiplantae;

REFERENCE
1 (bases 1 to 617)
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre,
P., McDermott, J. P., Shrager, D., Sillow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished (2001)

AUTHORS
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre,
P., McDermott, J. P., Shrager, D., Sillow, C. and Stern, D.

TITLE
Analyses of the Chlamydomonas reinhardtii Genome: A Model
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished (2001)

JOURNAL
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000

COMMENT
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
Location/Qualifiers

FEATURES
source
1. 617
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mc+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
), Lambda Zap II"
/note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr; see Wells et al., (2000). Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. Bluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."

BASE COUNT 114 a 205 c 204 g 93 t 1 others
ORIGIN

Query Match 25.5%; Score 612.8; DB 13; Length 617;
Best Local Similarity 99.5%; Pred. No. 4.3e-114;

Matches 614; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

608 CGAGCACTGAGGCGCGAGCGCGAGCGCGCGAGCGCGAGCGCGAGCTGCG 667
   |||||
1  CGAGCACTGAGGCGCGAGCGCGAGCGCGAGCGCGAGCGCGAGCGCGAGCTGCG 60
   |||||
668 CCGCGGCTGATGCTATGCTGAGAAATCTTACCGCGAGCTGATCCCTTACGTCAG 727
   |||||
61 CCGCGGCTGATGCTATGCTGAGAAATCTTACCGCGAGCTGATCCCTTACGTCAG 120
   |||||
728 CTGCAAGAGCCCGAGATGATGCTGCGCGCGAGCTGATGCTGATGCTGATGCTG 787
   |||||

```

Db	121	CTGCAGAGACCCCCAGATGATGTCGTCGGCCATGGCTCAAGTCTCTAGCGAAAGAA	180
Oy	788	GGGCATCCGGCCAAAGGACATGGTCTAATGGTGTCCATCATGACCCCTGACGCCGAACGATC	847
Db	181	GGGCACTCCGGCCAAAGGACATGGTCTAATGGTGTCCATCATGACCCCTGACGCCGAACGATC	240
Oy	848	GGAGGCTGACCGCGACTGAGTTCTGTGTGACCCGCCACCCCAACCTTGCGGCACAGTGA	907
Db	241	GGAGGCTGACCGCGACTGAGTTCTGTGTGACCCGCCACCCCAACCTTGCGGCACAGTGA	300
Oy	908	GCTCATCACCAACCGTGGAGCTGGGCAATCTTCAAGAGCGCGGGCATTAACCTGGCCGA	967
Db	301	GCTCATCACCAACCGTGGAGCTGGGCAACATCTTTCAAGAGCGCGGCATCAACCTGGCCGA	360
Oy	968	GCTGCCGAGGCGAGTGGGACAAATCCMAATGGAGCGTGGGCTCGGGGCGCGGCGTGTGT	102
Db	361	GCTGCCGAGGCGAGTGGGACAAATCCMAATGGAGCGTGGGCTCGGGGCGCGGCGTGTGT	420
Oy	1028	CGGCACCAACGGCGGTCATGTAGAGGCGGCGCTGCGCACGSCCTATAGCTTTCACGGG	108
Db	421	CGGCACCAACGGCGGTCATGTAGAGGCGGCGCTGCGCACGSCCTATAGCTTTCACGGG	480
Oy	1088	CACGCGCGCTGCGGCGCTGAGGCTGAGCGAGGTGGCGGSCATAGGACGGCATGAAGAGAC	114
Db	481	CACGCGCGCTGCGGCGCTGAGGCTGAGCGAGGTGGCGGSCATAGGACGGCATGAAGAGAC	540
Oy	1148	CAACATCAACATGTGTCCCGCGCCCGCGGTCCAAAGTTTGAAGAGCTGTGAACACCGCGC	120
Db	541	CAACATCAACATGTGTGTCCCGCGCCCGCGGTCCAAAGTTTGAAGAGCTGTGAACACCGGNC	600
Oy	1208	CGCCGCGCGCGCGGAGG 1224	
Db	601	CGTCGTGCGCGCGGAGG 617	

RESULT 6
BE724427
DOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BE724427 601 bp mRNA linear EST 14-SEP-2000
894076B10.y1 C. reinhardtii CC-1690, normalized, lambda Zap II
Chlamydomonas reinhardtii cDNA, mRNA sequence.
BE724427
BE724427.1 GI:10125723
EST.
Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 601)
Grossman, A., Davles, J., Federpiel, N., Harris, E., Jelebyre, P.,
McGermott, J. P., Siltfow, C., Stern, D. and Surzycki, R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 2
unpublished (2000)
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu
Location/Qualifiers
1..601

```

/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mc+ 21gr"
/db_xref="taxon:3055"
<clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in the
light, TAP medium in the dark, HS (minimal) medium in

```

ambient levels of CO₂ and HS medium bubbled with 5% CO₂. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBlueScript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

Query Match	24.9%; Score 596.4; DB 12; Length 601;
Best Local Similarity	99.5%; Pred. No. 9e-111;
Matches	597; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 17	CAAACTCTTCGACGCACTAGCTCTCAACCTCTGAAACCTTTTCCACAGTTTACACC 76
Db 2	CAAACTCTTCGACGCACTAGCTCTCAACCTCTGAAACCTTTTCCACAGTTTACACC 61
QY 77	CCAAATGGAGAGCGCGCTCCAAAGCTCGTCGCTCTCATATGACACACTATTTATT 136
Db 62	CCAAATGGAGAGCGCGCTTCACAGCTCGCTCCGTTGCTCTTATATGACACACTATTTATT 121
QY 137	CTAATATCTAGACGCGACAGATGTGCGGCTCTGTCCTGAAGCTTCGCGCGCGCTGTC 196
Db 122	CTAATATCTAGAGCGGACAAAGATGTGCGGCTCTGTCCTGAAGCTTCGCGCGCGCTGTC 181
QY 197	TATTGCGGCGACTCTTCGACGAGCGCGGACAGTGTGCCCCCGCGCTCCGCTCGACGACAG 256
Db 182	TATTGCGGCGCGACTCTTCGACGAGCGCGGACAGTGTGCCCCCGCGCTCCGCTCGACGACAG 241
QY 257	CACCGTGGGTAGACCTCTTGCAACACTTGAAGCGCGCCGACAGCGGCTAGGCAACGTTCGC 316
Db 242	CACCGTGGGTAGACCTCTTGCAACACTTGAAGCGCGCCGACAGCGGCTAGGCAACGTTCGC 301
QY 317	TTGCGCGGCTTCGACACCCGCTGCGGAGCGGCTTTGATGATGTCCAGCAGGCGCTTCGC 376
Db 302	TTGCGCGGCTTCGACACCCGCTGCGGAGCGGCTTTGATGATGTCCAGTAGGCGGCTTCGC 361
QY 377	CGACCTTGGCCAAAGCCCAAGGACGACCCACGCGCAACAGTCTGCGTGCAGAGTGGCTCC 436
Db 362	CGAGTTGGCCAAAGCCCAAGGACGACCCACGCGCAACAGTCTGCGTGCAGAGTGGCTCC 421
QY 437	GGCGGTTCTGTCTCTATTGCGGAGACCTTGCGGCTTCGCGCGCGGCGCCACACCCCAA 496
Db 422	GGCGGTTCTGTCTCTATTGCGGAGACCTTGCGGCTTCGCGCGCGGCGCGACACCCCAA 481
QY 497	GCACTGGCGGAGGGGCTTCGCGCGGCTTCGCTTTGACGAGGTGTTTGAACAGCTGTTTGG 556
Db 482	GCACTGGCGGAGGGGCTTCGCGCGGCTTCGCTTTGACGAGGTGTTTGAACAGCTGTTTGG 541
QY 557	CGCGGACCTGACATCATGAGAGGAGGAGGAGACTGTGTGACAGCGCTCAACCGAGCACT 616
Db 542	CGCGGACCTGACATCATGAGAGGAGGAGGAGACTGTGTGACAGCGCTTCAACCGAGCACT 601

RESULT 7	
AV389547	
LOCUS	558 bp mRNA linear EST 29-SEP-2000
DEFINITION	AV389547 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
ACCESSION	CDNA clone CM044d10_r, mRNA sequence.
VERSION	AV389547
KEYWORDS	AV389547.1 GI:6543763
SOURCE	EST.
ORGANISM	Chlamydomonas reinhardtii. Chlamydomonas reinhardtii. Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonadeae; Chlamydomonades. 1 (bases 1 to 558)
REFERENCE	Asamizu,E., Nakamura,Y., Sato,S., Fukunawa,H. and Tabata,S. A large scale structural analysis of cDNAs in a unicellular green alga, Chlamydomonas reinhardtii. I. Generation of 3433 non-redundant expressed sequence tags
AUTHORS	
TITLE	

JOURNAL MEDLINE
 20152988
 Contact: Yasukazu Nakamura
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: ynakamura@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 Location/Qualifiers
 1..558
 /organism="Chlamydomonas reinhardtii"
 /strain="C9"
 /db_xref="taxon:3055"
 /clone="CM044d10_r"
 /clone_1lb="Chlamydomonas reinhardtii C9"
 /dev_stage="photoautotrophic growth"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"
 BASE COUNT 113 a 171 c 202 g 72 t
 ORIGIN
 Query Match 23.2%; Score 556.4; DB 10; Length 558;
 Best Local Similarity 99.8%; Pred. No. 1.1e-102;
 Matches 557; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1213 CGCGCGCGGAGCGCGCGCGCGCGACACCCCGCGCGCTGGCTGGGACGGCGCGCG 1272
 Db 1 CGCGCGCGGAGCGCGCGCGCGCGACACCCCGCGCGCTGGCTGGGACGGCGCGCG 60
 Oy 1273 GCTTACCAGCGAGGACGGGCGGCGCATCACACTGCGGTCGCGCGCCGCGCAACGGGC 1332
 Db 61 GCTTACCAGCGAGGACGGGCGGCGCATCACACTGCGGTCGCGCGCCGCGCAACGGGC 120
 Oy 1333 TGGGCAAGCCCAAGAGCTGATCAACAGATGACGCGCGCGGAGGCCAAGTACACTTTG 1392
 Db 121 TGGGCAAGCCCAAGAGCTGATCAACAGATGACGCGCGCGGAGGCCAAGTACACTTTG 180
 Oy 1393 TGGAGATCATGGCCCTGCCCCCGGGCGTGTGTGGGCGGCGCGCGCCAGCCCGCTCCACG 1452
 Db 181 TGGAGATCATGGCCCTGCCCCCGGGCGTGTGTGGGCGGCGCGCGCCAGCCCGCTCCACG 240
 Oy 1453 ACAAGGCGCATCAACGAGAGCGCGGACGCGCGCTGTACAACTGTGACGAGAAGTCCACGC 1512
 Db 241 ACAAGGCGCATCAACGAGAGCGCGGACGCGCGCTGTACAACTGTGAGACAGAAGTCCACGC 300
 Oy 1513 TGGCGCGGACCCACGAGAACCCGTCATCCGCGGACCTGTACGACGACGTAACCTCGGAGAGC 1572
 Db 301 TGGCGCGGACCCACGAGAACCCGTCATCCGCGGACCTGTACGACGACGTAACCTCGGAGAGC 360
 Oy 1573 CGCTGGGCCCAAGAGCGCGACGAGCTGCTGACACACCCTACTGACCTGACCGCGCGCTGAGG 1632
 Db 361 CGCTGGGCCCAAGAGCGCGACGAGCTGCTGACACACCCTACTGACCTGACCGCGCGCTGAGG 420
 Oy 1633 AGAAGGACGAGAGAGTGAAGAGGCGCCACGAGGCTTTTGGGCGGAGACGTTCAAGC 1692
 Db 421 AGAAGGACGAGAGAGTGAAGAGGCGCCACGAGGCTTTTGGGCGGAGACGTTCAAGC 480
 Oy 1693 GAGGGGGCGATTAGAGAGTACCGTAAATATGCACTGATGGGTGATCGGGGTCTCCTCTT 1752
 Db 481 GAGGGGGCGATTAGAGAGTACCGTAAATATGCACTGATGGGTGATCGGGGTCTCCTCTT 540
 Oy 1753 TATATTGAATGGGTCAA 1770
 Db 541 TATATTGAATGGGTCAA 558
 RESULT 8
 LOCUS BG848385 602 bp mRNA linear EST 29-MAY-2001
 DEFINITION Chlamydomonas reinhardtii CC-1690, normalized, lambda Zap II
 ACCSSION BG848385
 VERSION BG848385.1 GI:14229569
 KEYWORDS EST.

SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT	FEATURES	
Chlamydomonas reinhardtii	Chlamydomonas reinhardtii	1 (bases 1 to 602)		Unpublished (2000)		source	
Chlamydomonas reinhardtii	Chlamydomonas reinhardtii	Grossman, A., Davies, J., Federle, N., Harris, E., Lefebvre, P., McDermott, J. P., Sillflow, C., Stern, D., and Surzycki, R.	Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unpublished (2000)	Contact: Charles Hauser			
Chlamydomonadaceae; Chlamydomonas			Vascular Plants; project phase 2	Contact: Charles Hauser			
				DCMB Box 91000			
				Duke University			
				Durham, NC 27708-1000			
				Tel: 919 613 8159			
				Fax: 919 613 8177			
				Email: chause@duke.edu			
				Location/Qualifiers			
				1. .602			
				/organism="Chlamydomonas reinhardtii"			
				/strain="CC-1690 wild type mcr+ 21gr"			
				/db_xref="taxon:3055"			
				/clone_1ib="CC. reinhardtii CC-1690, normalized, Lambda Zap II"			
				/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. POLYA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsist (Stratagene) phase. The library was normalized using method 4 described in Donald et al (1996) Genome Research 6: 791-806."			
BASE COUNT	134 a	141 c	184 g	143 t			
ORIGIN							
Query Match	22.7%	Score 545,	DB 12,	Length 602:			
Best Local Similarity	95.7%	Pred. No. 2.4e-100;					
Matches 560, Conservative	0;	Mismatches 25;	Indels 0;	Gaps 0;			
QY	1800	TGAGTGTCTCACAGCATGGGGACGTGTGGGGAGCGCAGTAGGCTGTTCACCTGACGCT	1859				
DB	18	TGCCCTGGCTCTTCCACATTTGATGTGACGGCGCTCCCATAGGCTGTTCACGACGCT	77				
QY	1860	GGCATTATGACGTAAGTATCTGGCATATGAGGAGACGCGGCTTGTCTAAACGATGCGATCC	1919				
DB	78	GGCATTATGACGTAAGTATCTGGCATATGAGGAGACGCGGCTTGTCTAAACGATGCGATCC	137				
QY	1920	TCACAGGACGTCGGAATGGCGCGTCCCATCATACGCAAAATTTCTGGGCTTCATGCTTC	1979				
DB	138	TCACAGGACGTCGGAATGGCGCGTCCCATCATACGCAAAATTTCTGGGCTTCATGCTTC	197				
QY	1980	TGGATTATGACGTCGCAAAACCTGCAATTTGCTTTGTTACAGTGGCCCAATCTT	2039				
DB	198	TGGATTATGACGTCGCAAAACCTGCAATTTGCTTTGTTACAGTGGCCCAATCTT	257				
QY	2040	GGTTGGAAGCTAAAGATGTTTGGGAACAATTCATCTTAATAAGCGTGTGGGGTTGAGG	2099				
DB	258	GGTTGGAAGCTAAAGATGTTTGGGAACAATTCATCTTAATAAGCGTGTGGGGTTGAGG	317				
QY	2100	ATGCGCAGTTGTGCGCTGTGTGGTGGCGGGGAACGTGGTGGCATTTAGCTTACCTGGC	2155				
DB	318	ATGCGCAGTTGTGCGCTGTGTGGTGGCGGGGAACGTGGTGGCATTTAGCTTACCTGGC	377				
QY	2160	ATACGACAAACGGGGCCCGTGAAGATTGAGCACTTGACCTGCGCACTTATGAACGTAGCGC	2211				
DB	378	ATACGACAAACGGGGCCCGTGAAGATTGAGCACTTGACCTGCGCACTTATGAACGTAGCGC	437				

QY 2220 TTATATCCACCGTATGCGATTGACCTTGGTGTAGGCAACGACCGGATAGGAAGCGGAG 2279
 DB 438 TTATATCCACCGTATGCGATTGACCTTGGTGTAGGCAACGACCGGATAGGAAGCGGAG 497
 QY 2280 AGATGATTCGAACCGCTGTAAAGAACGGCATAGTACTAGACATCTGTGATGAGACC 2339
 DB 498 AGATGATTCGAACCGCTGTAAAGAACGGCATAGTACTAGACATCTGTGATGAGACC 557
 QY 2340 CTGGGCGACGACGACGAGAGAGGTGTGATCAGCGCTGTAA 2384
 DB 558 CTGGGCGACGACGACGAGAGAGGTGTGATCAGCGCTGTAA 602

RESULT 9
 LOCUS B1716925
 DEFINITION 103101361.y1 C. reinhardtii CC-1690, Stress II (normalized),
 Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
 ACCESSION B1716925.1 GI:15692620
 VERSION B1716925.1
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii.
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

REFERENCE
 AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
 1 (bases 1 to 599)
 TITLE P., Mcdermott, J.P., Shrager, J., Silflow, C., and Stern, D.
 Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants. Project: 1031
 JOURNAL Unpublished (2001)
 COMMENT Contact: Charles Hauser
 DCMB Box 91000
 Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu.

FEATURES
 source location/Qualifiers
 1..599
 /organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
), Lambda Zap II"
 /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
 XhoI; Stress condition II library, constructed by John
 Davies and Jeffrey Mcdermott, combines cDNAs from CC-1690
 cells grown to mid-log phase in TAP (NH4+ - containing)
 and shifted to TAP - NO3- (24hrs); H2 production
 conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
 Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
 sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
 PolyA mRNA was purified from each sample, pooled and cDNA
 synthesized. The cDNA was directionally cloned into lambda
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
 sites. pBluescript II SK- plasmids were excised from the
 lambda Zap clones by superinfection with Exsistie
 (Stratagene) phage. The library was normalized using
 method 4 described in Bonaldi et al., (1996) Genome
 Research 6: 791-806."

BASE COUNT 123 a 172 c 207 g 97 t
 ORIGIN

Query Match 22.7%; Score 543.8; DB 13; Length 599;
 Best Local Similarity 99.6%; Pred. NO. 4.1e-100;
 Matches 545; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1298 CGGCATCACACTGCGCTGGCCGCGCAACGGCTGTGCAAGCAAGACCTGATCAC 1357
 DB 1 CGGCATCACACTGCGCTGGCCGCGCAACGGCTGTGCAAGCAAGACCTGATCAC 60

QY 1358 CAAGATGACAGCGCGGCGAGCCCAAGTACGACTTTGTGAGATATGCGCTGCCCGCGG 1417
 DB 61 CAAGATGACAGCGCGGCGAGCCCAAGTACGACTTTGTGAGATATGCGCTGCCCGCGG 120
 QY 1418 CTGTGTGGGCGGGGGGGGCGGCGGCGCTTCCACCGCAAGGCCATACCGCAAGCGGCA 1477
 DB 121 CTGTGTGGGCGGGGGGGGCGGCGGCGCTTCCACCGCAAGGCCATACCGCAAGCGGCA 180
 QY 1478 GGGCGGCTGTATCAACCTGGAGCAAGATGTCACGCTGGCGCGGCGCAAGAACCGCTG 1537
 DB 181 GGGCGGCTGTATCAACCTGGAGCAAGATGTCACGCTGGCGCGGCGCAAGAACCGCTG 240
 QY 1538 CATCCGCGAGCTGTACGACAGTACTCGAGAGCGCTGGCGCAAGGCGCAGAGCT 1597
 DB 241 CATCCGCGAGCTGTACGACAGTACTCGAGAGCGCTGGCGCAAGGCGCAGAGCT 300
 QY 1598 GCTGCACACCCACTTACGTCGCGCGCGCGCTGAGAGCAAGAGACGAGAGAGTGAAGAC 1657
 DB 301 GCTGCACACCCACTTACGTCGCGCGCGCGCTGAGAGCAAGAGAGAGTGAAGAC 360
 QY 1658 GCCAGAGGCTCTTGGGCGGAGACAGCTTCAAGCGAGGGGGGCTATTAGACGATCCGTA 1717
 DB 361 GCCAGAGGCTCTTGGGCGGAGACAGCTTCAAGCGAGGGGGGCTATTAGACGATCCGTA 420
 QY 1718 AATATGCACTGATGCGTGAATGCGGCTGCTCTCTTATATTGAATGGGCTCAAAATAGGC 1777
 DB 421 AATATGCACTGATGCGTGAATGCGGCTGCTCTCTTATATTGAATGGGCTCAAAATAGGC 480
 QY 1778 GGGGGGTCAAAATGTTTCTTTTGAATGCTGATCAACGATGGGCAAGTGTGCGAGGCC 1837
 DB 481 GGGGGGTCAAAATGTTTCTTTTGAATGCTGATCAACGATGGGCAAGTGTGCGAGGCC 540
 QY 1838 AGTAGGC 1844
 DB 541 AGTTGCC 547

RESULT 10
 LOCUS BG848384/c 621 bp mRNA linear EST 29-MAY-2001
 DEFINITION 1024021E05.x2 C. reinhardtii: CC-1690, normalized, Lambda Zap II
 Chlamydomonas reinhardtii cDNA, mRNA sequence.
 ACCESSION BG848384
 VERSION BG848384.1 GI:14229568
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii.
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonas reinhardtii
 1 (bases 1 to 621)
 TITLE Grossman, A., Davies, J., Federpiel, N., Harris, E., Lefebvre, P.,
 Mcdermott, J.P., Silflow, C., Stern, D., and Surzycki, R.
 Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants; project phase 2
 JOURNAL Unpublished (2000)
 COMMENT Contact: Charles Hauser
 DCMB Box 91000
 Duke University
 Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu.

FEATURES
 source location/Qualifiers
 1..621
 /organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
 II"
 /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
 XhoI; This library, constructed by John Davies and Jeffrey
 Mcdermott, combines cDNAs from CC-1690 cells, grown to

mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO₂ and HS medium bubbled with 5% CO₂. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 149 a 189 c 141 g 142 t

Query Match 22.5%; Score 539.8; DB 12; Length 621;
Best Local Similarity 93.8%; Pred. No. 2.7e-99;
Matches 562; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

1800 TGAAGTGTACAGCAGTGGGCGACGTCGGAGCGCAGTAGGCTGTCTACGACGCT 1859
604 TGCGTGTCTCTCCACATTTGTGTGAGCCGGCTCCGACATGGCTTTCACTGACGCT 545
1860 GGCATTAGCGCTAGTACTGACATGAGGAGCGCGCTTGTCTAACCGAATGCGTATCC 1919
544 GGCATTAGCGCTAGTACTGACATGAGGAGCGCGCTTGTCTAACCGAATGCGTATCC 485
1920 TCCAGGCGACGTCGGAAATGCGCGCTGCCATCAACGAAATCTTGCGCTTATCGCTTC 1979
484 TCCAGGCGACGTCGGAAATGCGCGCTGCCATCAACGAAATCTTGCGCTTATCGCTTC 425
1980 TGGATATTGAAGCTGCAACACCTGCAATTTATTTGTTTACAGTGTCCCAATCTT 2039
424 TGGATATTGAAGCTGCAACACCTGCAATTTATTTGTTTACAGTGTCCCAATCTT 365
2040 GGTGGAAGCTAAACATGTTTGGGAACAATTCATCTACTAAAGCGTGGGGTTAGG 2099
364 GGTGGAAGCTAAACATGTTTGGGAACAATTCATCTACTAAAGCGTGGGGTTAGG 305
2100 ATGCGACGTTGTGCGCTGTGGTGGGCGGAAAGTGGGAGCATTTAGGCTTACGTGC 2159
304 ATGCGACGTTGTGCGCTGTGGTGGGCGGAAAGTGGGAGCATTTAGGCTTACGTGC 245
2160 ATGCGAAGCGGGGCGCGTGAAGATTGACATTCGCGAATTTATTAAGCTTACGCGC 2219
244 ATGCGAAGCGGGGCGCGTGAAGATTGACATTCGCGAATTTATTAAGCTTACGCGC 185
2220 TTATATCCACCGTATCGATTGACGTTGTGTAGGACACGCGGTAGGAAGCGGAG 2279
184 TTATATCCACCGTATCGATTGACGTTGTGTAGGACACGCGGTAGGAAGCGGAG 125
2280 AGATGCTTTGCAACGCTGTAAAGAAAGCGCATAGCTACTAGACATCTGTGTGAC 2339
124 AGATGCTTTGCAACGCTGTAAAGAAAGCGCATAGCTACTAGACATCTGTGTGAC 65
2340 CTGGGCGACCGACGAGAGAGGTGTGATCGCGCTGTGTAGGACCGCATCTTGA 2398
64 CTGGGCGACCGACGAGAGAGGTGTGATCGCGCTGTGTAGGACCGCATCTTGA 6

RESULT 11
AV396217 538 bp mRNA linear EST 23-APR-2002
LOCUS AV396217 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
DEFINITION AV396217 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
ACCESSION AV396217
VERSION AV396217
KEYWORDS EST: GI:6550433
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 538)
REFERENCE
AUTHORS Aasamizu E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
TITLE A large scale structural analysis of cDNAs in a unicellular green

alga, Chlamydomonas reinhardtii. I. Generation of 3433 non-redundant expressed sequence tags
DNA Res. 6 (6), 369-373 (1999)

JOURNAL MEDLINE 20152988
COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/planc/.
Location/Qualifiers
1. 538
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="CL58B09_r"
/clone_lib="Chlamydomonas reinhardtii C9"
/dev_stage="photoautotrophic growth"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

FEATURES
source

BASE COUNT 83 a 202 c 153 g 100 t

Query Match 22.4%; Score 536.4; DB 10; Length 538;
Best Local Similarity 99.8%; Pred. No. 1.3e-98;
Matches 537; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

25 CTGCGAGGCACTAGCTCAAAACCTCGAACCCTTTTCCACAGTTTACACCCCAATTGC 84
1 CTGCGAGGCACTAGCTCAAAACCTCGAACCCTTTTCCACAGTTTACACCCCAATTGC 60
85 GACGCCCTTCGAAGCTGCTCCGTTGCTCTTATCATGACACCACTTATTTATATATC 144
61 GACGCCCTTCGAAGCTGCTCCGTTGCTCTTATCATGACACCACTTATTTATATATC 120
145 GTAGAGCGCAAGATGTGGGCGCTGTGTGTAAGCCCTGCGGCGGCTGTATTCGCG 204
121 GTAGAGCGCAAGATGTGGGCGCTGTGTGTAAGCCCTGCGGCGGCTGTATTCGCG 180
205 GCAGCTCTGCGAGGCGCGGCAAGTGCAGCCCGCGCTCCGCTCGCAGCGGACCGGTGC 264
181 GCAGCTCTGCGAGGCGCGGCAAGTGCAGCCCGCGCTCCGCTCGCAGCGGACCGGTGC 240
265 GTGTAGCCCTTGAACACTTGAAGCGCGCGGCAAGCGGCTGTATTCGCGG 324
241 GTGTAGCCCTTGAACACTTGAAGCGCGCGGCAAGCGGCTGTATTCGCGG 300
325 CTGCGCACCGCGTGTGCGAGGCGCTTGTGATGATGATGATGATGATGATGATGATG 384
301 CTGCGCACCGCGTGTGCGAGGCGCTTGTGATGATGATGATGATGATGATGATGATG 360
385 CCAAGCCCAAG 444
361 CCAAGCCCAAG 420
445 GTGTGCTATTTGCGAG 504
421 GTGTGCTATTTGCGAG 480
505 CCGAGGCGCTCGCGCGCTGCGCTTGTGACGAGGTGTTGACAGAGTGTGACAGAGTGTG 562
481 CCGAGGCGCTCGCGCGCTGCGCTTGTGACGAGGTGTTGACAGAGTGTGACAGAGTGTG 538

RESULT 12
AV396237 499 bp mRNA linear EST 23-APR-2002
LOCUS AV396237 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
DEFINITION AV396237 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
ACCESSION AV396237
VERSION AV396237
KEYWORDS EST: GI:6550453
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 499)
Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
A large scale structural analysis of cDNAs in a unicellular green
alga, Chlamydomonas reinhardtii. I. Generation of 3433
non-redundant expressed sequence tags

JOURNAL
MEDLINE
COMMENT
20152988
Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
Location/Qualifiers

FEATURES

source

1..499
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="CL58a12_r"
/clone_1ib="Chlamydomonas reinhardtii C9"
/dev_stage="photoautotrophic growth"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT
ORIGIN
77 a 191 c 141 g 90 t

Query Match
Best Local Similarity 100.0%; Score 499; DB 10; Length 499;
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 CTCGAACCTCGAATCTTTTCACAGATTACACCCCAATTCGAGCCGCTCCACGC 99
Db 1 CTCGAACCTCGAATCTTTTCACAGATTACACCCCAATTCGAGCCGCTCCACGC 60
QY 100 TCGCTCCGTTGCTCTTCAATCGCAACCACTATTATTCTTAATTCGTAGACGGACA 159
Db 61 TCGCTCCGTTGCTCTTCAATCGCAACCACTATTATTCTTAATTCGTAGACGGACA 120
QY 160 TGTCCGCGCTCGTCTGTAAGCCCTGCGCGCGCTGTCTAATTCGCGGACACTCTCTGAGGG 219
Db 121 TGTCCGCGCTCGTCTGTAAGCCCTGCGCGCGCTGTCTAATTCGCGGACACTCTCTGAGGG 180
QY 220 CCGGAGAGGTCCGCGCGCGCTCCGCTCGACGACGACCGGTGATAGCCCTTGCA 279
Db 181 CCGGAGAGGTCCGCGCGCGCTCCGCTCGACGACGACCGGTGATAGCCCTTGCA 240
QY 280 CACTTAGAGGCGCGGACGCGCGCTTAGGCAACGTGCTTGCGGGCTGCGGACCCGCTG 339
Db 241 CACTTAGAGGCGCGGACGCGCGCTTAGGCAACGTGCTTGCGGGCTGCGGACCCGCTG 300
QY 340 CGAGAGCGCTTGAATCATGTCCAGACAGCGCTCCGCCAGCTTGCCAAAGCCCAAGAG 399
Db 301 CGAGAGCGCTTGAATCATGTCCAGACAGCGCTCCGCCAGCTTGCCAAAGCCCAAGAG 360
QY 400 ACCCCACGCGACGACAGTCTGCTGAGGTGCTCGGCGCTTGCTGATGCTATTATGCG 459
Db 361 ACCCCACGCGACGACAGTCTGCTGAGGTGCTCGGCGCTTGCTGATGCTATTATGCG 420
QY 460 AGACCTTGAGGCTTGAGGCGCGGCGGCGACCAACCCCAAGACGCTGCGGAGGGCTTCGCG 519
Db 421 AGACCTTGAGGCTTGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 520 GCTCGGCTTTGACGAGGT 538
Db 481 GCTCGGCTTTGACGAGGT 499

RESULT 13
AV386674 498 bp mRNA linear EST 29-SEP-2000
LOCUS
DEFINITION
AV386674 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
cDNA clone CM004a11_r, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Chlamydomonas reinhardtii.
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 498)
Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
A large scale structural analysis of cDNAs in a unicellular green
alga, Chlamydomonas reinhardtii. I. Generation of 3433
non-redundant expressed sequence tags

JOURNAL
MEDLINE
COMMENT
20152988
Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
Location/Qualifiers

FEATURES

source

1..498
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="CM004a11_r"
/clone_1ib="Chlamydomonas reinhardtii C9"
/dev_stage="photoautotrophic growth"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT
ORIGIN
78 a 194 c 139 g 87 t

Query Match
Best Local Similarity 100.0%; Score 498; DB 10; Length 498;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 CTCGAGGCACTAGCTTCAACCTCGAATCTTTTCACAGATTACACCCCAATTCG 84
Db 1 CTCGAGGCACTAGCTTCAACCTCGAATCTTTTCACAGATTACACCCCAATTCG 60
QY 85 GACGCGCTCCAGCTCGCTCGCTGCTCTTCAATCGCAACCACTATTATTCTTAATTC 144
Db 61 GACGCGCTCCAGCTCGCTCGCTGCTCTTCAATCGCAACCACTATTATTCTTAATTC 120
QY 145 GTAGACGAGCAAGATGTGCGGCTGTGCTGAACCCCTGCGGCGGCGGTCTATTTCGG 204
Db 121 GTAGACGAGCAAGATGTGCGGCTGTGCTGAACCCCTGCGGCGGCGGTCTATTTCGG 180
QY 205 GCAGCTCTGACGAGGCGCGGACAGTGCAGCCCGCGCTTCGCTCGACGACACCGTGC 264
Db 181 GCAGCTCTGACGAGGCGCGGACAGTGCAGCCCGCGCTTCGCTCGACGACACCGTGC 240
QY 265 GTGTAGCCCTTGCAACACTTGAAGGCGCGGACGCGCGCTTAGGCAACGTGCTTGCGCG 324
Db 241 GTGTAGCCCTTGCAACACTTGAAGGCGCGGACGCGCGCTTAGGCAACGTGCTTGCGCG 300
QY 325 CTGCGGCAACCGCTGCGGAGGCGCTTGTAGTCATGTCCAGAGGCGCTGCGGAGCTTG 384
Db 301 CTGCGGCAACCGCTGCGGAGGCGCTTGTAGTCATGTCCAGAGGCGCTGCGGAGCTTG 360
QY 385 CCAAGCCCAAGAGACACCCACGCGCAAGACAGTGTGCTGAGGTGCTTCGCGCGTTC 444
Db 361 CCAAGCCCAAGAGACACCCACGCGCAAGACAGTGTGCTGAGGTGCTTCGCGCGTTC 420
QY 445 GTGTGCTATTGCTCCAGACCTTGGGCTTGGCGCGGCGGCGGCGGCGGCGGCGGCGG 504
Db 421 GTGTGCTATTGCTCCAGACCTTGGGCTTGGCGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 505 CCGAGGCGCTCGCGCGCC 522
Db 481 CCGAGGCGCTCGCGCGCC 498

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
BI1717903/c	550 bp mRNA linear EST 19-SEP-2001	BI1717903	1	GI:15693598	EST.	Chlamydomonas reinhardtii.	Chlamydomonas reinhardtii.	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.	1 (bases 1 to 550)	Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shager, J., Silflow, C. and Stern, D.	Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1031	Unpublished (2001)
DCMB Box 91000	Duke University	Durham, NC 27708-1000	Tel: 919 613 8159	Fax: 919 613 8177	Email: chauuser@duke.edu.	Location/Qualifiers	1..550	/organism="Chlamydomonas reinhardtii"	/strain="CC-1690 wild type mt+ 21gr"	/db_xref="taxon:3055"	/clone_11b="C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II"	/note="Vector: pBluescript II SK-; Site, 1: EcoRI, Site, 2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Wells et al., (2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into Lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the Lambda Zap clones by superinfection with ExBst18 (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 793-806."
BASE COUNT	131 a	167 c	126 g	126 t								
ORIGIN												
Query Match	19.9%;	Score 477.6;	DB 13;	Length 550;								
Best Local Similarity	95.3%;	Pred. No. 1.1e-86;										
Matches 492;	Conservative 0;	Mismatches 24;	Indels 0;	Gaps 0;								
QY	1800	TGAGTGTGTACAGCATGGGCGACGTGTGCGGAGCCAGTGGCTTCACTGACGCT	1855	TGAGTGTGTACAGCATGGGCGACGTGTGCGGAGCCAGTGGCTTCACTGACGCT								
DB	522	TGCGCTGCTCTTCCACATTGGGTAGGCGCGCTCCGCATAGGCTGTTCACATGACGT	463	TGCGCTGCTCTTCCACATTGGGTAGGCGCGCTCCGCATAGGCTGTTCACATGACGT								
QY	1860	GGCATTAGGCGTACTGTGCATGTGAGGAGCGCGCTTGCTTAACCGAATGGCGTATCC	1915	GGCATTAGGCGTACTGTGCATGTGAGGAGCGCGCTTGCTTGATGCGGCGGCTTCTT								
DB	462	GGCATTAGGCGTACTGTGCATGTGAGGAGCGCGCTTGCTTAACCGAATGGCGTATCC	403	GGCATTAGGCGTACTGTGCATGTGAGGAGCGCGCTTGCTTGATGCGGCGGCTTCTT								
QY	1920	TCCAGGCGACGTGCGAATGGCGCGTCCCATCAACGCAATCTTGCGCTTATGCGCTTC	1975	TCCAGGCGACGTGCGAATGGCGCGTCCCATCAACGCAATCTTGCGCTTATGCGCTTC								
DB	402	TCCAGGCGACGTGCGAATGGCGCGTCCCATCAACGCAATCTTGCGCTTATGCGCTTC	343	TCCAGGCGACGTGCGAATGGCGCGTCCCATCAACGCAATCTTGCGCTTATGCGCTTC								
QY	1980	TGCAATATTGAAGCTGCACAAACCTGCATTTATTTGCTTTTACACGTGCCCAATCTT	2033	TGCAATATTGAAGCTGCACAAACCTGCATTTATTTGCTTTTACACGTGCCCAATCTT								
DB	342	TGCAATATTGAAGCTGCACAAACCTGCATTTATTTGCTTTTACACGTGCCCAATCTT	283	TGCAATATTGAAGCTGCACAAACCTGCATTTATTTGCTTTTACACGTGCCCAATCTT								
QY	2040	GGTTGGAACTGAACATGTTGGGAACAATTCATCTTACTTAAGCGTGGGGGCTTGAG	2095	GGTTGGAACTGAACATGTTGGGAACAATTCATCTTACTTAAGCGTGGGGGCTTGAG								

Accession	Sequence	Position
Db	282 GGTGGAGCGCTAAACATGTTGGGAACAATTCACTTACTAAACGCTGTGGGGGTGAGG	223
Oy	2100 ATGGCAGCATGTGTCGCTGCTGTGGGTGGCGGGAACGTGGGTAGCATTTTAGCTAGTGGC	2159
Db	222 ATGGGCACGTTGTGCGCTGGTGGGTGGGCGGGAACGTGGGTAGCATTTAGGCTAGCTGGC	163
Oy	2160 ATACGACAAACGGGGCCCGTAGAGATTAGACACTTGACCTGACCGGAACTTATGAACGTAGCGC	2219
Db	162 ATACGACAAACGGGGCCCGTAGAGATTAGACCTTGACTCCGGAACTTATGAACGTAGCGC	103
Oy	2220 TTATATACCAACCGTATGGAATTGACGTTGTGTATGAGCAACGACGCGTAGGAAGCGGAG	2275
Db	102 TTATATACCCCGCATGCGATTGACGTTGTGTATGAGCAACGACGCGGTAGGAAGCGGAG	43
Oy	2280 AGATGCAATTGCAAAACCGCTGTAAAGAAAGAACGACATAG	2315
Db	42 AGATGCAATTGCAAAACCGCTGTAAAGAAAGAACGACATAG	7

BI532214/c	BI532214	631 bp	MRNA	linear	EST 29-AUG-2001
LOCUS	1041119G32.xl C. reinhardtii CC-1690, normalized,	Lambda Zap II			
DEFINITION	Chlamydomonas reinhardtii cDNA, mRNA sequence.				
ACCESSION	BI532214				
VERSION	BI532214.1	GI:15372788			
KEYWORDS	EST.				
SOURCE	Chlamydomonas reinhardtii.				
ORGANISM	Chlamydomonas reinhardtii				
	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;				
	Chlamydomonadaceae; Chlamydomonas.				
REFERENCE	1 (bases 1 to 631)				
AUTHORS	Grosman, A., Chang, C.-W., Davies, J.J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shrager, J., Stillflow, C. and Stern, D.				
TITLE	Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1024b				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Charles Hauser DCMB Box 91000 Duke University Durham, NC 27708-1000 Tel.: 919 613 8159 Fax: 919 613 8177 Email: chauser@duke.edu.				
FEATURES	Location/Qualifiers				
source	1..631				
	/organism="Chlamydomonas reinhardtii"				
	/strain="CC-1690 wild type mt+ 21gr"				
	/db_xref="taxon:3055"				
	/clone_lib="C. reinhardtii CC-1690, normalized, Lambda zap				
	II"				
	/note="Vector: pBluescript II SK-, Site 1: EcoRI, Site 2:				
	XhoI; This library, constructed by John Davies and Jeffrey				
	McDermott, combines cDNAs from CC-1690 cells grown to				
	mid-log phase in TAP (acetate-containing) medium in the				
	light, TAP medium in the dark, HS (minimal) medium in				
	ambient levels of CO2 and HS medium bubbled with 5% CO2.				
	POLYA mRNA was purified from each sample, pooled and cDNA				
	synthesized. The cDNA was directionally cloned into lambda				
	ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.				
	pBluescript II SK- plasmids were excised from the lambda				
	ZAP clones by superinfection with Exsistit (Stratagene)				
	phage. The library was normalized using method 4 described				
	in Bonaldo et al (1996) Genome Research 6: 791-806."				
BASE COUNT	148 a 195 c 147 g 141 t				
ORIGIN					
Query Match	19.8%; Score 476; DB 13; Length 631;				
Best Local Similarity	95.2%; Pred.No. 2,3e-86;				
Matches 491; Conservative	0; Mismatches 25; Indels 0; Gaps 0;				

```
QY 1800 TGAGTGTGTCAACAGATGGGGACCTGTGCGGAGGCCAGTAGCTGTTCATCTGCACGCT 1859
Db 522 TGCCTGTCTCTTCCACATTGTGAGCGCGCTCCGCATAGCTGTTCATCTGCACGCT 463
QY 1860 GGCATTAGGCGTAGTACTGTGCATGAGGGAGCGCGCTTGCTAACCGAATGGCGTATCCC 1919
Db 462 GGCATTAGGCGTAGTACTGTGCATGAGGGAGCGCGCTTGCTAACCGAATGGCGTATCCC 403
QY 1920 TCCAGGGGACGTGGAATGGCGCGTGCATCAACGCAAAATCTTGCCCTTCATCGCTTC 1979
Db 402 TCCAGGGGACGTGGAATGGCGCGTGCATCAACGCAAAATCTTGCCCTTCATCGCTTC 343
QY 1980 TGGATATTGAAGCTGCACAAACCTGCATTCTATTGCTTGTACAGCTGCCCAATCTT 2039
Db 342 TGGATATTGAAGCTGCACAAACCTGCATTCTATTGCTTGTACAGCTGCCCAATCTT 283
QY 2040 GGTGGAAGCTTAACATGTTTGGGAACATTCATCTTAAGCGGTGGGGGTGAGG 2099
Db 282 GGTGGAAGCTTAACATGTTTGGGAACATTCATCTTAAGCGGTGGGGGTGAGG 223
QY 2100 ATGCGCACGTTGTGCGCTGTGGTGGGCGGGAACGTGGGTAGCATTAGGCTAGCTGGC 2159
Db 222 ATGCGCACGTTGTGCGCTGTGGTGGGCGGGAACGTGGGTAGCATTAGGCTAGCTGGC 163
QY 2160 ATACGACAAAGGGGCCGTGAGGATTGAGCACTTGACTGCGCACTTATGAACGTAGCGC 2219
Db 162 ATACGACAAAGGGGCCGTGAGGATTGAGCACTTGACTGCGCACTTATGAACGTAGCGC 103
QY 2220 TTTATACCCACCGTATGCGATTGACGTTGTGTAGGCAACGAGCCGTAGGAAGCGGAG 2279
Db 102 TTTATACCCACCGTATGCGATTGACGTTGTGTAGGCAACGAGCCGTAGGAAGCGGAG 43
QY 2280 AGATGCATTGCAAAACGCTGTAAAGAACGCGCATAG 2315
Db 42 AGATGCATTGCAAAACGCTGTAAAGAACGCGCATAG 7
```

Search completed: June 8, 2003, 12:43:28
Job time : 3037 secs

This Page Blank (uspto)